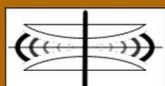


International Journal on Advances in Systems and Measurements



The *International Journal on Advances in Systems and Measurements* is published by IARIA.

ISSN: 1942-261x

journals site: <http://www.ariajournals.org>

contact: petre@aria.org

Responsibility for the contents rests upon the authors and not upon IARIA, nor on IARIA volunteers, staff, or contractors.

IARIA is the owner of the publication and of editorial aspects. IARIA reserves the right to update the content for quality improvements.

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy or print, providing the reference is mentioned and that the resulting material is made available at no cost.

Reference should mention:

International Journal on Advances in Systems and Measurements, issn 1942-261x
vol. 11, no. 1 & 2, year 2018, http://www.ariajournals.org/systems_and_measurements/

The copyright for each included paper belongs to the authors. Republishing of same material, by authors or persons or organizations, is not allowed. Reprint rights can be granted by IARIA or by the authors, and must include proper reference.

Reference to an article in the journal is as follows:

<Author list>, "<Article title>"
International Journal on Advances in Systems and Measurements, issn 1942-261x
vol. 11, no. 1 & 2, year 2018, http://www.ariajournals.org/systems_and_measurements/

IARIA journals are made available for free, proving the appropriate references are made when their content is used.

Sponsored by IARIA

www.aria.org

Copyright © 2018 IARIA

Editors-in-Chief

Constantin Paleologu, University "Politehnica" of Bucharest, Romania
Sergey Y. Yurish, IFSA, Spain

Editorial Advisory Board

Vladimir Privman, Clarkson University - Potsdam, USA
Winston Seah, Victoria University of Wellington, New Zealand
Mohammed Rajabali Nejad, Universiteit Twente, the Netherlands
Nageswara Rao, Oak Ridge National Laboratory, USA
Roberto Sebastian Legaspi, Transdisciplinary Research Integration Center | Research Organization of Information and System, Japan
Victor Ovchinnikov, Aalto University, Finland
Claus-Peter Rückemann, Westfälische Wilhelms-Universität Münster / Leibniz Universität Hannover / North-German Supercomputing Alliance, Germany
Teresa Restivo, University of Porto, Portugal
Stefan Rass, Universität Klagenfurt, Austria
Candid Reig, University of Valencia, Spain
Qingsong Xu, University of Macau, Macau, China
Paulo Esteveao Cruvinel, Embrapa Instrumentation Centre - São Carlos, Brazil
Javad Foroughi, University of Wollongong, Australia
Andrea Baruzzo, University of Udine / Interaction Design Solution (IDS), Italy
Cristina Seceleanu, Mälardalen University, Sweden
Wolfgang Leister, Norsk Regnesentral (Norwegian Computing Center), Norway

Indexing Liaison Chair

Teresa Restivo, University of Porto, Portugal

Editorial Board

Jemal Abawajy, Deakin University, Australia
Ermeson Andrade, Universidade Federal de Pernambuco (UFPE), Brazil
Francisco Arcega, Universidad Zaragoza, Spain
Tulin Atmaca, Telecom SudParis, France
Lubomír Bakule, Institute of Information Theory and Automation of the ASCR, Czech Republic
Andrea Baruzzo, University of Udine / Interaction Design Solution (IDS), Italy
Nicolas Belanger, Eurocopter Group, France
Lotfi Bendaouia, ETIS-ENSEA, France
Partha Bhattacharyya, Bengal Engineering and Science University, India
Karabi Biswas, Indian Institute of Technology - Kharagpur, India

Jonathan Blackledge, Dublin Institute of Technology, UK
Dario Bottazzi, Laboratori Guglielmo Marconi, Italy
Diletta Romana Cacciagrano, University of Camerino, Italy
Javier Calpe, Analog Devices and University of Valencia, Spain
Jaime Calvo-Gallego, University of Salamanca, Spain
Maria-Dolores Cano Baños, Universidad Politécnica de Cartagena, Spain
Juan-Vicente Capella-Hernández, Universitat Politècnica de València, Spain
Vítor Carvalho, Minho University & IPCA, Portugal
Irinela Chilibon, National Institute of Research and Development for Optoelectronics, Romania
Soolyeon Cho, North Carolina State University, USA
Hugo Coll Ferri, Polytechnic University of Valencia, Spain
Denis Collange, Orange Labs, France
Noelia Correia, Universidade do Algarve, Portugal
Pierre-Jean Cottinet, INSA de Lyon - LGEF, France
Paulo Esteveao Cruvinel, Embrapa Instrumentation Centre - São Carlos, Brazil
Marc Daumas, University of Perpignan, France
Jianguo Ding, University of Luxembourg, Luxembourg
António Dourado, University of Coimbra, Portugal
Daniela Dragomirescu, LAAS-CNRS / University of Toulouse, France
Matthew Dunlop, Virginia Tech, USA
Mohamed Eltoweissy, Pacific Northwest National Laboratory / Virginia Tech, USA
Paulo Felisberto, LARSyS, University of Algarve, Portugal
Javad Foroughi, University of Wollongong, Australia
Miguel Franklin de Castro, Federal University of Ceará, Brazil
Mounir Gaidi, Centre de Recherches et des Technologies de l'Energie (CRTE), Tunisie
Eva Gescheidtova, Brno University of Technology, Czech Republic
Tejas R. Gandhi, Virtua Health-Marlton, USA
Teodor Ghetiu, University of York, UK
Franca Giannini, IMATI - Consiglio Nazionale delle Ricerche - Genova, Italy
Gonçalo Gomes, Nokia Siemens Networks, Portugal
Luis Gomes, Universidade Nova Lisboa, Portugal
Antonio Luis Gomes Valente, University of Trás-os-Montes and Alto Douro, Portugal
Diego Gonzalez Aguilera, University of Salamanca - Avila, Spain
Genady Grabarnik, CUNY - New York, USA
Craig Grimes, Nanjing University of Technology, PR China
Stefanos Gritzalis, University of the Aegean, Greece
Richard Gunstone, Bournemouth University, UK
Jianlin Guo, Mitsubishi Electric Research Laboratories, USA
Mohammad Hammoudeh, Manchester Metropolitan University, UK
Petr Hanáček, Brno University of Technology, Czech Republic
Go Hasegawa, Osaka University, Japan
Henning Heuer, Fraunhofer Institut Zerströrungsfreie Prüfverfahren (FhG-IZFP-D), Germany
Paloma R. Horche, Universidad Politécnica de Madrid, Spain
Vincent Huang, Ericsson Research, Sweden
Friedrich Hülsmann, Gottfried Wilhelm Leibniz Bibliothek - Hannover, Germany
Travis Humble, Oak Ridge National Laboratory, USA

Florentin Ipate, University of Pitesti, Romania
Imad Jawhar, United Arab Emirates University, UAE
Terje Jensen, Telenor Group Industrial Development, Norway
Liudi Jiang, University of Southampton, UK
Kenneth B. Kent, University of New Brunswick, Canada
Fotis Kerasiotis, University of Patras, Greece
Andrei Khrennikov, Linnaeus University, Sweden
Alexander Klaus, Fraunhofer Institute for Experimental Software Engineering (IESE), Germany
Andrew Kusiak, The University of Iowa, USA
Vladimir Laukhin, Institució Catalana de Recerca i Estudis Avançats (ICREA) / Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Spain
Kevin Lee, Murdoch University, Australia
Wolfgang Leister, Norsk Regnesentral (Norwegian Computing Center), Norway
Andreas Löf, University of Waikato, New Zealand
Jerzy P. Lukaszewicz, Nicholas Copernicus University - Torun, Poland
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Sathiamoorthy Manoharan, University of Auckland, New Zealand
Stefano Mariani, Politecnico di Milano, Italy
Paulo Martins Pedro, Chaminade University, USA / Unicamp, Brazil
Don McNickle, University of Canterbury, New Zealand
Mahmoud Meribout, The Petroleum Institute - Abu Dhabi, UAE
Luca Mesin, Politecnico di Torino, Italy
Marco Mevius, HTWG Konstanz, Germany
Marek Miskowicz, AGH University of Science and Technology, Poland
Jean-Henry Morin, University of Geneva, Switzerland
Fabrice Mourlin, Paris 12th University, France
Adrian Muscat, University of Malta, Malta
Mahmuda Naznin, Bangladesh University of Engineering and Technology, Bangladesh
George Oikonomou, University of Bristol, UK
Arnaldo S. R. Oliveira, Universidade de Aveiro-DETI / Instituto de Telecomunicações, Portugal
Aida Omerovic, SINTEF ICT, Norway
Victor Ovchinnikov, Aalto University, Finland
Telhat Özdoğan, Recep Tayyip Erdogan University, Turkey
Gurkan Ozhan, Middle East Technical University, Turkey
Constantin Paleologu, University Politehnica of Bucharest, Romania
Matteo G A Paris, Università degli Studi di Milano, Italy
Vittorio M.N. Passaro, Politecnico di Bari, Italy
Giuseppe Patanè, CNR-IMATI, Italy
Marek Penhaker, VSB- Technical University of Ostrava, Czech Republic
Juho Perälä, Bitfactor Oy, Finland
Florian Pinel, T.J.Watson Research Center, IBM, USA
Ana-Catalina Plesa, German Aerospace Center, Germany
Miodrag Potkonjak, University of California - Los Angeles, USA
Alessandro Pozzebon, University of Siena, Italy
Vladimir Privman, Clarkson University, USA
Mohammed Rajabali Nejad, Universiteit Twente, the Netherlands

Konandur Rajanna, Indian Institute of Science, India
Nageswara Rao, Oak Ridge National Laboratory, USA
Stefan Rass, Universität Klagenfurt, Austria
Candid Reig, University of Valencia, Spain
Teresa Restivo, University of Porto, Portugal
Leon Reznik, Rochester Institute of Technology, USA
Gerasimos Rigatos, Harper-Adams University College, UK
Luis Roa Oppliger, Universidad de Concepción, Chile
Ivan Rodero, Rutgers University - Piscataway, USA
Lorenzo Rubio Arjona, Universitat Politècnica de València, Spain
Claus-Peter Rückemann, Leibniz Universität Hannover / Westfälische Wilhelms-Universität Münster / North-German Supercomputing Alliance, Germany
Subhash Saini, NASA, USA
Mikko Sallinen, University of Oulu, Finland
Christian Schanes, Vienna University of Technology, Austria
Rainer Schönbein, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Germany
Cristina Seceleanu, Mälardalen University, Sweden
Guodong Shao, National Institute of Standards and Technology (NIST), USA
Dongwan Shin, New Mexico Tech, USA
Larisa Shwartz, T.J. Watson Research Center, IBM, USA
Simone Silvestri, University of Rome "La Sapienza", Italy
Diglio A. Simoni, RTI International, USA
Radosveta Sokullu, Ege University, Turkey
Junho Song, Sunnybrook Health Science Centre - Toronto, Canada
Leonel Sousa, INESC-ID/IST, TU-Lisbon, Portugal
Arvind K. Srivastav, NanoSonix Inc., USA
Grigore Stamatescu, University Politehnica of Bucharest, Romania
Raluca-Ioana Stefan-van Staden, National Institute of Research for Electrochemistry and Condensed Matter, Romania
Pavel Šteffan, Brno University of Technology, Czech Republic
Chelakara S. Subramanian, Florida Institute of Technology, USA
Sofiene Tahar, Concordia University, Canada
Muhammad Tariq, Waseda University, Japan
Roald Taymanov, D.I.Mendeleyev Institute for Metrology, St.Petersburg, Russia
Francesco Tiezzi, IMT Institute for Advanced Studies Lucca, Italy
Wilfried Uhring, University of Strasbourg // CNRS, France
Guillaume Valadon, French Network and Information and Security Agency, France
Eloisa Vargiu, Barcelona Digital - Barcelona, Spain
Miroslav Velev, Aries Design Automation, USA
Dario Vieira, EFREI, France
Stephen White, University of Huddersfield, UK
Shengnan Wu, American Airlines, USA
Qingsong Xu, University of Macau, Macau, China
Xiaodong Xu, Beijing University of Posts & Telecommunications, China
Ravi M. Yadahalli, PES Institute of Technology and Management, India
Yanyan (Linda) Yang, University of Portsmouth, UK

Shigeru Yamashita, Ritsumeikan University, Japan

Patrick Meumeu Yomsi, INRIA Nancy-Grand Est, France

Alberto Yúfera, Centro Nacional de Microelectronica (CNM-CSIC) - Sevilla, Spain

Sergey Y. Yurish, IFSA, Spain

David Zammit-Mangion, University of Malta, Malta

Guigen Zhang, Clemson University, USA

Weiping Zhang, Shanghai Jiao Tong University, P. R. China

CONTENTS

pages: 1 - 12

A Comprehensive Framework for High Resolution Image-based 3D Modeling and Documentation of Crime Scenes and Disaster Sites

Sven Becker, University of Applied Sciences Mittweida, Germany
Michael Spranger, University of Applied Sciences Mittweida, Germany
Florian Heinke, University of Applied Sciences Mittweida, Germany
Steffen Grunert, University of Applied Sciences Mittweida, Germany
Dirk Labudde, University of Applied Sciences Mittweida, Germany

pages: 13 - 21

MobileTimeSync - An Android App for Time Synchronization for Mobile Construction Assessment

Maik Benndorf, University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Germany
Marika Kaden, University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Germany
Eric Zuchantke, Friedrich-Schiller-Universität Jena, Faculty of Mathematics and Computer Science, Germany
Frederic Ringsleben, University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Germany
Thomas Haenselmann, University of Applied Sciences Mittweida, Faculty of Applied Computer Sciences & Biosciences, Germany
Martin Gaedke, Technische Universität Chemnitz, Faculty of Computer Science, Germany

pages: 22 - 35

Evaluating Key Factors Influencing ERTMS Risk Assessment: a Reference Model

Katja Schuitemaker, University of Twente, Netherlands
Heidi van Spaandonk, ProRail, The Netherlands
Marco Kuijsten, NS, The Netherlands
Mohammad Rajabalinejad, University of Twente, The Netherlands

pages: 36 - 46

Univariate Modeling of the Timings and Costs of Unknown Future Project Streams: A Case Study

Alireza Shojaei, University of Florida, United States
Ian Flood, University of Florida, United States

pages: 47 - 60

Progressive Advancement of Knowledge Resources and Mining: Integrating Content Factor and Comparative Analysis Methods for Dynamical Classification and Concordances

Claus-Peter Rückemann, Westfälische Wilhelms-Universität Münster (WWU) and Leibniz Universität Hannover and HLRN, Germany

pages: 61 - 71

Modeling of Gas Compressors and Hierarchical Reduction for Globally Convergent Stationary Network Solvers

Tanja Clees, Fraunhofer Institute for Algorithms and Scientific Computing, Germany
Igor Nikitin, Fraunhofer Institute for Algorithms and Scientific Computing, Germany
Lialia Nikitina, Fraunhofer Institute for Algorithms and Scientific Computing, Germany
Lukasz Segiet, Fraunhofer Institute for Algorithms and Scientific Computing, Germany

pages: 72 - 82

Scientific Challenges in Archaeology - Is modern computer science ready for archaeology?

Lutz Schubert, University of Ulm, Germany

Gill Hunt, Hunt Lancaster, United Kingdom

Keith Jeffery, Keith G Jeffery Consultants, United Kingdom

pages: 83 - 92

QoS-Ensured Cooperative Vehicular Communications

Radwa Ahmed Osman, Arab Academy for Science, Technology & Maritime Transport Alexandria, Egypt

Xiao-Hong Peng, Aston University, United Kingdom

pages: 93 - 99

Measurement System Based on Arduino for Biogas Sensing: Development Considerations and Laboratory Scale Approach

Isaías González Pérez, University of Extremadura, España

Antonio José Calderón Godoy, University of Extremadura, España

pages: 100 - 110

Low Power Optimized and DPA Resistant D-FF for Versatile Mobile Applications

Karol Niewiadomski, Chair of Automation and Computer Science, University of Wuppertal, Germany

Dietmar Tutsch, Chair of Automation and Computer Science, University of Wuppertal, Germany

pages: 111 - 123

Capillary Sensors with UV-Forced Degradation and Fluorescence Reading of Chemical Stability and Polycyclic Aromatic Hydrocarbons Presence in Diesel Fuels

Michał Borecki, Institute of Microelectronics and Optoelectronics, Warsaw University of Technology, Poland

Michael L Korwin -Pawlowski, Département d'informatique et d'ingénierie, Université du Québec en Outaouais, Canada

Mateusz Gęca, Institute of Microelectronics and Optoelectronics, Warsaw University of Technology, Poland

Przemysław Prus, Blue Oak Inventions, Poland

Jan Szmidt, Institute of Microelectronics and Optoelectronics, Warsaw University of Technology, Poland

pages: 124 - 136

Investigating Different Ballistic Threats to Validate the Simulation Model of Fiber-Reinforced Plastics - Experimental Model Validation of Fiber Composites Under Ballistic Impact

Arash Ramezani, University of the Federal Armed Forces Hamburg, Germany

Hendrik Rothe, University of the Federal Armed Forces Hamburg, Germany

pages: 137 - 152

A Framework for Optimizing Simulation Model Validation & Verification

Bill Roungas, Delft University of Technology, The Netherlands

Sebastiaan Meijer, KTH Royal Institute of Technology, Sweden

Alexander Verbraeck, Delft University of Technology, The Netherlands

pages: 153 - 165

Numerical Investigation of the Attenuation of Shock Waves by Simulating a Second, Transient Medium to Protect Vehicles Against Blasts

Burghard Hillig, University of the Federal Armed Forces Hamburg, Germany

Arash Ramezani, University of the Federal Armed Forces Hamburg, Germany

Hendrik Rothe, University of the Federal Armed Forces Hamburg, Germany

pages: 166 - 172

Advances in the Development of Combat Helmet Systems - Constructional Contributions for Lightweight Ballistic Composites

Henrik Seeber, Helmut Schmidt University - University of the Federal Armed Forces, Germany
Arash Ramezani, Helmut Schmidt University - University of the Federal Armed Forces, Germany
Hendrik Rothe, Helmut Schmidt University - University of the Federal Armed Forces, Germany

pages: 173 - 182

Modeling and Simulation of Enzyme-Catalysed Substrate Conversion in a Microbioreactor

Linus Petkevičius, Institute of Computer Science, Vilnius University, Lithuania
Romas Baronas, Institute of Computer Science, Vilnius University, Lithuania

pages: 183 - 195

GeSCo: Exploring the Edge Beneath the Cloud in Decentralized Manufacturing

Badarinath Katti, TU Kaiserslautern, Kaiserslautern, Germany
Christiane Plociennik, DFKI GmbH, Kaiserslautern, Germany
Michael Plociennik, SAP SE, Walldorf, Germany

pages: 196 - 207

Fake Reviews Detection on Movie Reviews through Sentiment Analysis Using Supervised Learning Techniques

Elshrif Elmurngi, École de Technologie Supérieure, Canada
Abdelouahed Gherbi, École de Technologie Supérieure, Canada

pages: 208 - 219

Edge Computing and Blockchains for Flexible and Programmable Analytics in Industrial Automation

Mauro Isaja, Engineering Ingegneria Informatica SpA, Italy
John Soldatos, Athens Information Technology, Greece
Nikos Kefalakis, Athens Information Technology, Greece
Volkan Gezer, DFKI, Germany

pages: 220 - 229

Assessing the Application of the In-Memory Technology - A Comprehensive Framework

Stephan Ulbricht, Friedrich Schiller University Jena, Germany
Marek Opuszko, Friedrich Schiller University Jena, Germany
Johannes Ruhland, Friedrich Schiller University Jena, Germany

A Comprehensive Framework for High Resolution Image-based 3D Modeling and Documentation of Crime Scenes and Disaster Sites

Sven Becker*, Michael Spranger*, Florian Heinke*, Steffen Grunert* and Dirk Labudde*[†]

*University of Applied Sciences Mittweida, Germany

Faculty Applied Computer Sciences & Biosciences

Email: *{name.surname}@hs-mittweida.de* [†]Fraunhofer

Cybersecurity

Darmstadt, Germany

Email: labudde@hs-mittweida.de

Abstract—Serious crime scenes or disaster sites with many victims after natural disasters, airplane crashes or terrorist attacks require extensive and comprehensive investigations to clarify all circumstances leading to the event, to identify the victims and to find the responsible people. The results of investigations do not only serve to prosecute the perpetrators, yet are mainly used to develop preventive strategies and to strengthen the resilience of the society. 3D reconstructions of crime scenes and disaster sites provide a quick and comprehensive method to support the work of investigators digitally. The detailed reconstruction of the scene allows not only a long term documentation, yet also the simulation of alternative scenarios. To create a 3D reconstruction a large number of parameters has to be taken into account in order to make the necessary decisions. To this point there is no standardized procedure for the reconstruction of serious crime scenes or disaster sites. The framework presented in this paper can serve as a simple guide to create a 3D reconstruction. However, it can also be easily implemented in an already existing forensic process chain of investigative services. In this sense, it can form the basis for a standardization and thus ensure the comparability of different models. Additionally, a detailed description of an application of this framework for a real crime scene is given.

Keywords—forensic; resilience engineering; framework; 3D reconstruction; photogrammetry

I. INTRODUCTION

A. Background

Many forensic issues require the reconstruction of crime scenes or disaster sites in order to allow a quick and detailed investigation. In addition to identifying victims, perpetrators and third parties involved, it is of particular interest to gain knowledge in order to develop preventative strategies in the context of resilience engineering. Furthermore, information obtained from ad hoc digital models can help to control the situation, especially in rough terrain, by providing a tool to support the coordination of emergency services. Using the example of the Germanwings aircraft crash in March 2015

in the French Alps, it was shown how aerial photogrammetry can be used to create 3D models to manage the situation and elucidating such major damage events [1]. Another area of application is the reconstruction of crime scenes and subsequent simulation of possible courses of events documented in way that can be presented in front of a court of law. Previously used methods are based on laser scanning technology, as shown in [2]–[8]. Disadvantages of these technologies are high acquisition costs as well as limited mobility. As a result the use of these technologies is only reserved for a few special forces. However, photo equipment and smartphones have become a nationwide standard for all emergency services. It is therefore obvious to use these devices for the reconstruction of crime scenes and allows an analytical 3D examination of each crime scene and event site.

B. Literature Review

In photogrammetry, an alternative to laser scanners, images taken with a camera are used to create 3D models. With a software these images are aligned and the corresponding camera positions are calculated. In a next step, 3D point clouds are created, which are then used to create a 3D model. For taking these pictures several camera systems are available, yet most commonly SLR cameras are used. Even though these cameras are used for photogrammetry in several areas, for example in the documentation of archeological sites [9], underwater research [10], forensics like analysis of road accidents [11], [12] and the documentation of autopsies [13], of particular interest for this paper are 3D reconstructions of larger outdoor areas and buildings based on pictures taken with SLRs. Already in 2004 Kersten et al. showed the potential of images taken with a single-lens reflex camera to create 3D models using photogrammetry [14]. At that time, they used a combination of automated and other methods, to conduct a reconstruction of a castle. In a later study they once again used images of buildings taken with an SLR camera, which were processed using open-source, proprietary and web-based software solutions. The results were compared to laser scans and

the authors found that SLR images processed with open-source software can deliver 3D models of comparable quality [15], [16]. Similar results were reported by Falkingham et al., who took photographs of archaeological objects with an Olympus Megapixel Camera and processed them with appropriate open-source software [17]. Furthermore, Ziegler et al. were also able to positively evaluate the use of SLR cameras in combination with open source software for 3D modeling in their work from 2014 [18]. Additionally, several research groups analyzed the suitability of SLR images as well as the use of photogrammetric software solutions to calculate 3D models. Kersten et al. and Falkingham et al. focused on terrestrial images, yet also on how pictures taken from a helicopter or UAV may be used as an additional source. Actually, pictures taken with the help of helicopters or UAVs, for example drones, play an essential role in the research on photogrammetric 3D models for example of disaster sites. In 2008 Püschel et al. showed the potential of taking images of the Landenberg castle using drones with SLR cameras as payload [8]. Moreover, in the review of UAVs for photogrammetry and remote sensing by Colomina et al. it becomes clear to the reader, that there is a great advantage of using drones and their corresponding payloads for the creation of 3D models of large disaster sites [19]. Furthermore, Urbanova et al. discussed in their study the use of drone-mounted cameras for the documentation of bodies and show possible applications for the forensic sectors, especially the digitalization of crime scenes. Additionally, in 2015 Resig showed that in addition to drones, other flying devices, for example a kite, can be used to collect aerial images used for 3D models [20]. However, for the reconstruction of large disaster sites, not only buildings, yet also larger areas play a crucial role. Bendea et al, Naidoo et al. and Barazzetti et al. discuss in their studies the successful application of the aforementioned methodologies for the reconstruction of disaster sites, accidents and crime scenes as well as search and rescue operations [21]–[23]. Even though the aforementioned studies all made great contributions to the research of photogrammetry for the creation of 3D models, they only focus on methodologies, techniques and technologies and do not provide a consistent approach starting with the planning process up until finishing the 3D model. Hence, Kersten et al. present a straightforward framework for creating 3D models using photogrammetry, which can be used to plan a scene reconstruction, yet unfortunately more information about how to proceed beyond the planning process is not given [15]. Again, Püschel et al. specify in their study how collected data is processed, yet keep their descriptions very general and do not provide more detailed information [8]. Zancajo-Blazquez et al. present a comprehensive workflow for the 3D reconstruction of crime scenes, but focus in their studies rather on interiors than on larger areas [24]. Similarly, in their framework Kim et al. and González-Aguilera et al. describe the 3D reconstruction of disaster sites and larger areas as well as crime scenes based on images and open-source software, yet do not give information about what parameters are available in the software applications, which makes it difficult to derive a consistent approach [25]–[27]. Even though Gindraux et al. give a basic approach for the software Agisoft PhotoScan in their study from 2017,

including recommendations for settings of software-specific parameters, it is only useful for a quick reconstruction and if the user has enough knowledge about the equipment and the image and procedural requirements. Therefore, there is no universal description for the 3D reconstruction of crime scenes and disaster sites, which gives sufficiently detailed information about all necessary process steps and parameters. Frameworks that provide an overview of necessary process steps from the planning phase to the dissemination of results are already to be found in areas such as cultural inheritance as well as intelligence and reconnaissance. Makantasis et al. present a framework for the 3D documentation and reconstruction of cultural monuments from data acquisition to the processed model. They describe possible calculation errors of the 3D models and approaches to optimize the results in detail. Unfortunately, they provide only general information about the remaining process steps [28]. Xu et al. describe in their framework the 3D reconstruction of a cultural monument based on video recordings. In addition to the description of data acquisition using UAVs, information about the data processing are given by means of software applications such as VisualSfM in combination with PMVS. The framework provides a helpful overview of necessary process steps, but requires expert knowledge in many places [29]. Napolitano et al. presented a framework for the conservation of cultural heritage sites and provide additional information about incurred costs and the duration of individual process steps such as data collection. Despite the necessity of this information, unfortunately, there are only few details given about for example equipment for planning or parameterization for data processing [30]. Torres-Martínez et al. describe a comprehensive framework in the field of reconnaissance and remote sensing from the planning phase up to visualization possibilities of finished 3D models. Despite the clear and comprehensive information in some places only a few details are given such as necessary parametrization for processing collected data. Furthermore a lot of expert knowledge is required. However, this framework provides a basic overview of necessary process steps [31].

C. Objectives

To the knowledge of the authors up to this date, there is no consistent approach on 3D-reconstruction in the context of forensic investigations. Therefore, a comprehensive framework addressing this issue is presented in this paper. It relies on previous research and applications in the field of photogrammetric modeling, as described in the previous section. In Section II-A the entire framework is presented followed by a detailed description of all modules and sub-steps. Section III shows the reconstruction of one test scenario and one cold case as an example of how this framework can be implemented in practice. In addition, the influence of various parameters on model quality and computation time is demonstrated. The work concludes with a brief summary and gives an outlook on further work.

II. COMPREHENSIVE FRAMEWORK FOR CRIME SCENE AND DISASTER SITE RECONSTRUCTION

A. General Framework

The process of reconstructing forensically relevant items and complex crime scenes or disaster sites is similar, which is why the same schema can be used for all of them. The process usually consists of four to six modules (Fig. 1), which are selected depending on the purpose of the reconstruction and what data is used.

- **Planning:** operational coordination, strategy and resource planning as well as technology choice.
- **Data Collection:** parameterization, data collection, data preprocessing to ensure a uniform and consistent data base as a requirement for the following steps.
- **VSFM – CMPMVS – MeshLab/Agisoft Processing:** generation of 3D models
- **Data Integration:** integration of detailed models to supplement and complete a specific professional perspective (for example biological traces).
- **Simulation:** simulation of courses of events allow the mutual exclusion of alternative courses as well as the development of prevention strategies.
- **Documentation:** filing and court-proof documentation of all relevant model parameters, source data and results to ensure traceability of procedures and results

The starting point is always an incident or an inquiry by the prosecution, the defense or a judge. Subsequently, in the planning module decisions have to be made regarding what resources, technologies and strategies should be employed. Afterwards, within the data collection module, different ways are presented to collect the data and prepare it, after setting important parameters, for the creation of the 3D model. At this point, the decision has to be made whether to create the 3D model using open source or proprietary software. This decision is affected by factors like cost, training time, functionality, flexibility, portability as well as limitations in the image format. As a result of the respective processing module, a basic 3D model is generated. If the scene to be reconstructed as a 3D model is a more complex scene such as a crime scene, it has to be decided whether additional items, for example evidence, people or objects, should be added iteratively as 3D models. Consequently, the reconstruction of every item again starts with the planning module. Is the 3D model such an item it can either be integrated in an already existing scene or it can be documented as a separate 3D model. If no additional items have to be added, a simulation of the course of events can be done, if required. The result of the simulation is a realistic, complex, physically correct and fully parameterizable 3D model with integrated animations showing the course of events. The collected data in the form of images, the methodical procedures and techniques, which were used, the software-specific parameterizations as well as the calculated 3D models and simulations have to be documented in a way they can be used before a court of law. This can be

achieved by following the steps of the documentation module. In the following section a detailed description of each module will be given.

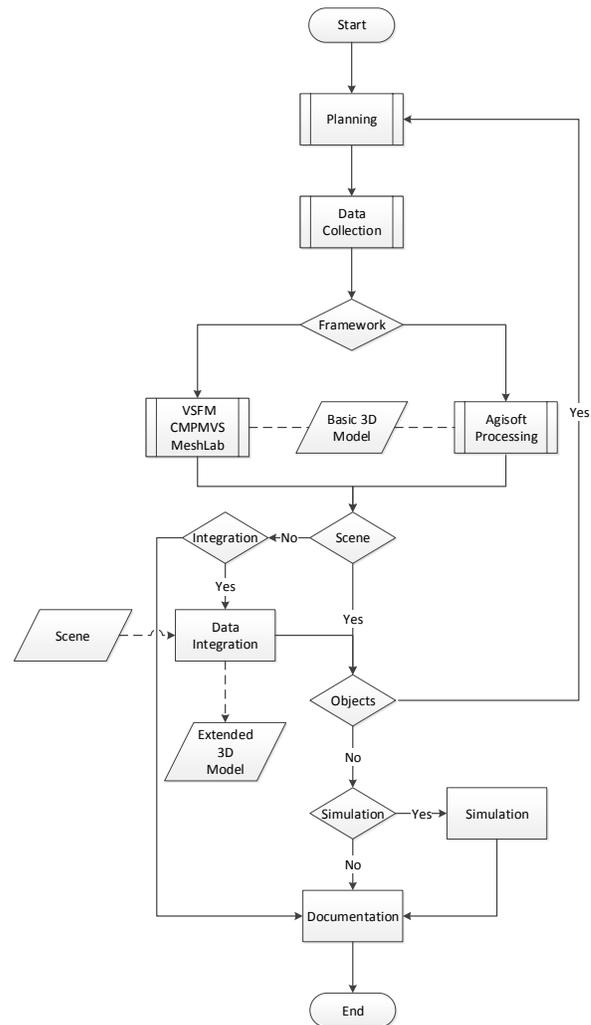


Figure 1. Framework for photogrammetric 3D reconstruction of crime scenes and disaster sites.

B. Planning Module

The “Planning Module” as illustrated in Fig. 2 is constituted of three general tasks:

- 1) the description of the scenario,
- 2) selection of an appropriate strategy, and
- 3) the orchestration of necessary equipment.

At the beginning, it is important to describe the item that is going to get reconstructed. The item can be a single object as part of a complex scene or the scene itself. For the latter it is

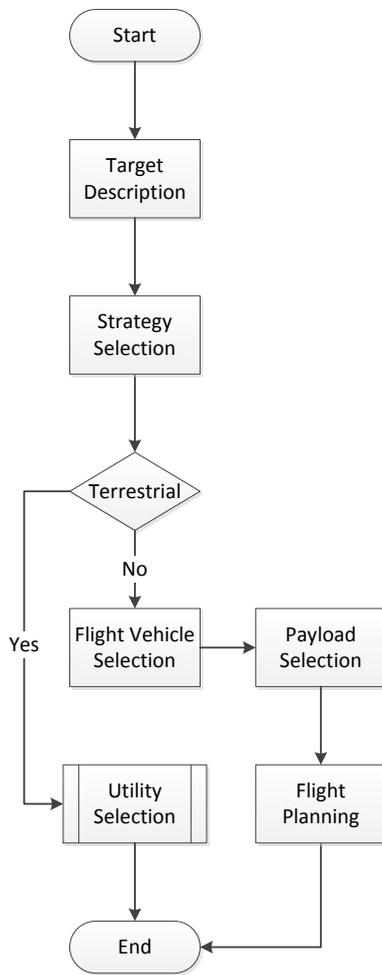


Figure 2. Description of the Planning Module.

important to determine, whether the site to be reconstructed is indoors or outdoors and other aspects that might influence the process of collecting the data and the reconstruction, including vegetation, buildings and possible obstacles like high-voltage lines. Furthermore, it is important to assess the accessibility of the site in general. With the help of the detailed description it is possible to choose a suitable strategy for taking the pictures or making the video. Three different strategies are possible: linear, contour-aligned or circular (Fig. 3). In most cases taking the pictures or making the video in a circular manner is the best option, especially for small items or aerial images as long as the terrain allows to do so and the flight path is not affected by any obstacles. The linear technique is especially suitable for the reconstruction of walls (inside or outside), facades and the reconstruction of long stretches of ground for example streets or rivers. However, when reconstructing buildings the contour-aligned method has many advantages,

especially for buildings with a contorted layout. After choosing the right technique, the decision has to be made, whether the data should be collected from the ground or with the help of a flying device. If the data is collected from the ground one can choose without further preparations, which hardware to use such as photography/video equipment (like camera, objective and filters) or an automatic rotator for taking pictures of single objects. Experiments showed when using an automatic rotator a homogenously colored wall in the background in combination with a stationary camera ensures a 3D object of high quality. If aerial images/videos are required, at first the flying device has to be chosen, for example a helicopter or a fitting drone type. Afterwards, one can choose the respective payload (e.g., camera, thermal sensors or gas sensors) depending on the loading capacity of the flying device and the data that needs to be collected. Finally, taking into account the description, the chosen strategy to collect the data and the chosen flying device, one can start to plan the flight using a suitable software (manufacturer-specific or open access sources like GoogleMaps) in consideration of the environmental conditions.

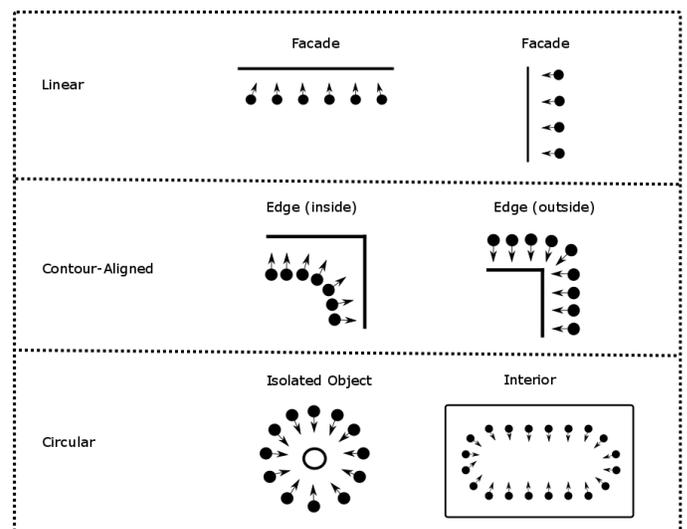


Figure 3. Overview about various strategies for data collection.

C. Data Collection Module

The process of collecting data as illustrated in Fig. 4 is one of the most important steps during the reconstruction process, since the decisions made here, for example about the parameterization of the hardware used, directly influence the quality of the generated 3D models.

- 1) Type: selection of the type of recording (video and images),
- 2) Parameter setting: parameterization of hardware used,
- 3) Collection: actual data acquisition,
- 4) Pre-Processing: conversion to an appropriate file format for further processing steps

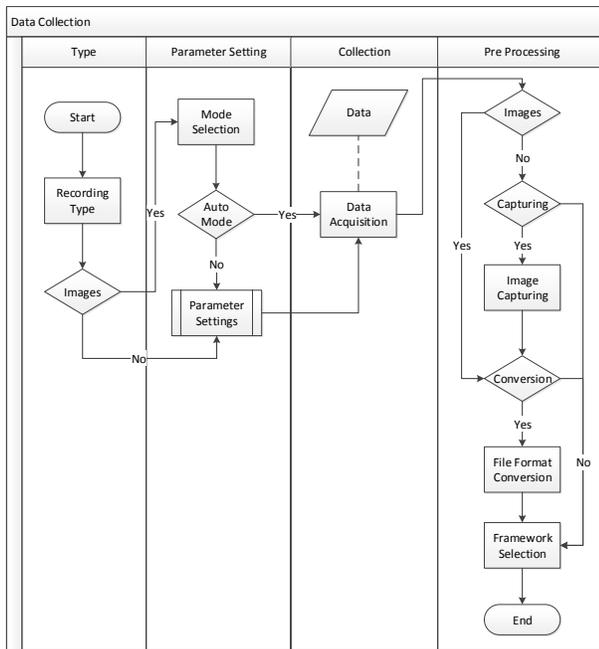


Figure 4. Overview of the data collection module.

Depending on the device used, the image or video file format must be chosen. It is recommended to use video recordings, because the acquisition of data is much quicker, and to only use photographs if time is not a concern. When taking a sequence of pictures, one has to decide, which image recording mode is most suitable (for example day/night, landscape, macro, car etc.). In case of using the automatic mode, the picture taking can be started without any further adjustments. For video recordings or when using a mode other than automatic, depending on the hardware it may be necessary to set up other parameters, such as the aperture or the shutter speed, before the actual picture taking or video making can be carried out. Furthermore, it is recommended to set the ISO value as low as possible to prevent additional noise in the images and, correspondingly, use a high f-number for a sufficient depth of field. Regardless of the mode and the selected parameters, it is recommended to use the highest camera resolution that is possible (full HD resolution for videos). Additionally, for objects it is recommended to use a mode that allows taking continued shots in order to have images with a sufficient overlap. After collected the data it must be decided, which model generation software (open source or Agisoft PhotoScan) should be used. However, in the case of video recordings, single images must be extracted beforehand. Depending on what software was chosen it has to be decided whether any preprocessing steps are necessary, such as filtering, enhancing the image quality or converting it to another image file format.

D. 3D Disaster-Site Reconstruction

1) *VSFM – CPMVS – MeshLab*: Open source software applications are a low-cost option to create 3D objects or complex scenarios based on images. The process to create models using photogrammetry as presented in Fig. 5 is based on the applications *VisualSFM*, *CPMVS* and *MeshLab* and consists of the following two general tasks:

- 1) Structure from Motion: 3D model generation of single objects or complex scenes, and
- 2) Post Processing: editing and optimization of generated models.

The first step in creating a 3D model is to calculate a point cloud of the aligned and overlapping photographs. For the calculation of the point cloud one can use the software package *VisualSFM*, which is mentioned in several papers [32] [33] and was used in previous work by the authors [1] [34]. However, before calculating the point cloud necessary parameters have to be set such as the maximal number of calculated features or the maximal acceptable resolution of the photographs used. Subsequently, the process step *Execute SIFT* determines reference points in the pictures, which are saved as key point descriptors and are the basis for creating the point cloud images. Nonetheless, for the actual *3D Scene Reconstruction* a sparse point cloud, based on a minimal number of features will be calculated first. This sparse point cloud can then be used to automatically create a textured model using the software *CPMVS*, which was mentioned in several papers [35] [18]. The process consist of the following three steps:

- 1) Execute CMVS (Clustering Views for Multi-view Stereo): fragmentation of the sparse point cloud into single clusters for quicker processing,
- 2) Execute PMVS (Patch-based Multi-view Stereo Software): enrichment of the sparse point cloud and calculation of a dense point cloud, and
- 3) Texturing: creation of a textured mesh

An alternative would be the sub steps CMVS und PMVS, also included as a function in *VisualSFM*. In this case the resulting model would be a dense cloud, which would have to be transformed into a textured mesh with the software *MeshLab* [36]. Afterwards, the mesh can be used in the subsequent process steps. However, our own experiments have shown that it is difficult to create good models with this method, because of the enormous parametrization. The calculated 3D model can be optimized in *MeshLab* (possible output formats: *.ply, *.stl, *.off, *.obj, *.3ds, *.vrm1 2.0, *.u3d, *.x3d and *.dae). Because of insufficient quality of the pictures, it is often necessary to edit individual areas of the model, for example by deleting artefacts or filling in holes.

2) *Agisoft PhotoScan*: An alternative to an open-source software application is the software Agisoft PhotoScan, a proprietary software for creating 3D models using photographs or videos. The process of creating a 3D model is illustrated in Fig. 6 and can be divided into two tasks:

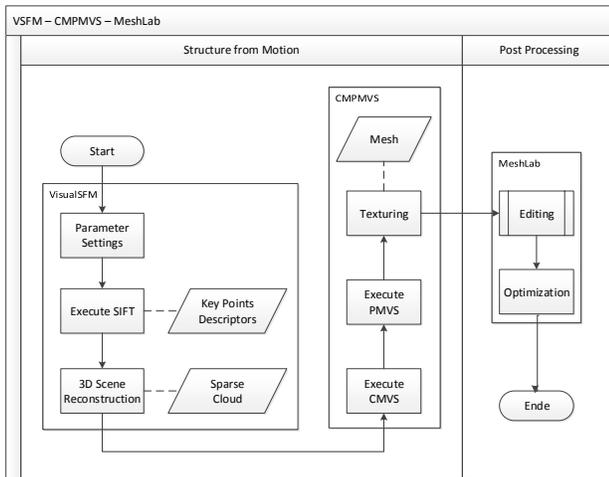


Figure 5. Detailed VSFM – CMPMVS – MeshLab processing module.

- 1) Pre-Processing: check image quality and camera calibration,
- 2) Model Generation: generation of fully textured model

After importing the photographs, they are displayed as thumbnails, which allow a visual examination of the photographs regarding their completeness and quality. An advantage of Agisoft PhotoScan is the option to quantify the quality of the photographs in a range of $[0 - 1]$ and the manufacturer recommends to exclude all pictures with an estimated quality of 0.5 or less. While the photographs are imported the software will automatically check by means of the EXIF data, whether a camera calibration – the compensation of the distortion of the lens – has already taken place. If the compensation is missing, it can be calculated mathematically by entering manually the lens parameters such as the radial distortion coefficient, the distortion center and the tangential distortion coefficient. After setting some parameters that influence the quality, the sparse point cloud is calculated using a similar concept as described in section II-D1. After some further optimization steps, for example removing redundant points or eliminating projection errors, the sparse point cloud is densified into a dense point cloud. Before the texture is transferred in a final step from the photographs onto the model, the points have to be interconnected to create a 3D mesh (possible output formats: *.obj, *.3ds, *.3ds, *.dae, *.ply, *.stl, *.dxf, *.fbx, *.u3d, *.wrl, *.kmz, *.pdf).

E. Data Integration

In many cases, it is necessary to integrate additional objects in the generated 3D scenes. In particular, for the simulation of possible courses of events, 3D models of, for example, evidence, victims or objects used in the commission of a crime are often essential and indispensable. In general, small objects can be added to larger scenes using the software already

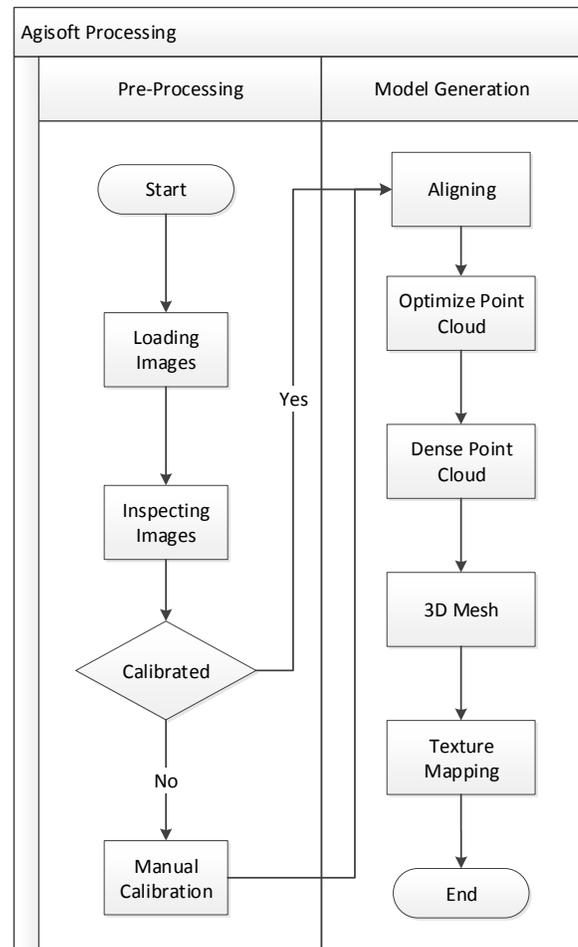


Figure 6. Detailed Agisoft processing module.

mentioned in the sections II-D1 and II-D2. However, in view of an eventual simulation it is recommendable to integrate the objects with the simulation software Blender. The simulation of the course of events can be started as soon as all necessary objects are integrated in the extended 3D model. If no additional objects are needed, one can skip this part and start directly with the documentation.

F. Simulation and Documentation

For the analysis of possible circumstances of a crime the reconstructed 3D scenes can be used to carry out computer-aided simulations. The simulations can aid in comparing alternative interactions between people (e.g., perpetrator and victim), between people and objects (e.g., weapon and victim) or between several objects (e.g., a projectile and a wall) and thus may help in drawing conclusions about the actual course of events. An additional application is the simulation of the

spread of biological or chemical pollutants in accidents or investigations of environmental crimes. For the computer-aided simulations, a variety of software applications is available. At this point it is recommended to use the open-source modeling software “Blender”. The integrated physics engine allows to visualize complex physical processes in form of three-dimensional simulations. After creating single objects, complex 3D scenarios and simulations, all the steps taken including all decisions made, what values for the parameters were chosen and, which software applications were used as well as the results, must be documented in a way they can be presented in front of a court of law.

III. PROTOTYPICAL APPLICATION

In this section two scenarios demonstrating the utilization of the proposed framework are discussed. In addition a brief summary on technical details of used hardware is given. The first scenario involves the employment of various imaging techniques to the water tower located in Mittweida and its surrounding area, which provides an ideal test ground due to the size of the area and the tower as well as small details aimed to be captured accurately in the reconstruction process. The second scenario is a real-life application to a crime scene in Jena Lobeda. In discussing both scenarios, more details on program usage and parameterization are given, thus extending the elucidations on the proposed frameworks towards their utilization in practice.

A. Technical Details

1) *Drone*: The drone used in our study is a *MikroKopter MK-ARF OktoXL 6S12*, an eight-blade rotary wing drone for multi-purpose utilization. In our set-up the MK OktoXL 6S12 has a maximum slant range of 4,000 m and a maximum ceiling of 5,000 m above sea level. With fully charged batteries and optimal weather conditions, the drone achieves a maximum flight time of about 30 minutes. Besides present weather conditions, maximum flight time is reduced by hardware additionally mounted to the drone, such as a fixed SLR camera. Furthermore, the drone is equipped with a CMOS camera whose video feed can be received and post-processed on the ground. Although not-movable around the yaw axes, both the CMOS camera and SLR camera mount can be pitched and rolled. In combination with automated pre-planned waypoint flight and POI focusing capabilities as well as automated camera triggering, the user is thus able to obtain images made in-flight at pre-planned positions and altitudes. Waypoints and trigger events can be set-up and uploaded to the drone using the maintenance and control software *MikroKopter-Tool* (software version used in this study: V2.12a).

2) *Camera Equipment*: If the used drone systems do not have a built-in camera, as it is the case for the MK-ARF OktoXL 6S12, additional camera equipment is needed. In principle, any mountable camera can be used, but factors such as resolution, parameterization of images and data format should be considered. The systems used should support a minimum

resolution of 1920x1080 pixels, allow a parameterization of aperture, shutter speed, ISO and other parameters, as well as support data formats such as RAW or JPG. The camera used in this study is a Nikon D7100 SLR camera. Due to the compact design and the low weight of 765 grams, this camera is ideal as a payload for the presented drone system. The Nikon D7100 features an optical low pass filter and 24.1 megapixel CMOS sensor. There are 51 measuring fields available for the autofocus system. The camera allows continuous shooting at 6 frames per second with a maximum possible resolution of 6,000 x 4,000 pixels and video recordings with a maximum of 1,920 x 1,080 pixels. Supported file formats are RAW and JPG.

3) *Motor-driven Turntable*: For capturing smaller objects such as ensured traces, the use of an automatic, motor-driven turntable is recommended. In this study, a model of the company *stageonair* was used, which allows a 360° object imagery. The diameter of the turntable is 62.5 cm, the load is designed for 150 kg. The associated software allows individual adjustment of the direction of rotation and speed of the turntable as well as the number of images, which can be realized using the automatic camera release. If the possibility using such systems is not given, due to its practicability in field utilization, lighter but not motorized turntables can be used as an alternative.

B. Test Scenario I — Mittweida Water Tower

The area in question is well-suitable for testing the proposed framework. The tower is about 38 m high and has a base diameter of 10 to 16 m. The surrounding area (about 150 m in diameter) is rather flat with no obstructing obstacles present. Small details (< 50 cm) on the tower surface provide references for qualitative evaluation of resolution and accuracy of obtained models. In Figure 7 models of the area are shown. Both models are based on images obtained by a drone-mounted SLR camera. As proposed, the drone was programmed to fly a circular flight path (50 m radius) at constant height above ground level. In this case an altitude of 50 m was chosen in order to provide a 45° look-down camera angle. Images were extracted from recorded HD video material every single second, resulting to 111 images with a resolution of 1,920 x 1,080 pixels, and processed as proposed. Although major details are discernible, smaller features with a size of about less than one meter are difficult to identify. On a standard desktop machine (eight 3.5 GHz CPUs, 32GB RAM, GeForce 750 GTX Titan), the 3D reconstruction process required about 1.5 to 2 hours of computation time and 1.5GB of disk space using the Open-Source software framework, about 2 to 2.5 hours and 2GB of disk space using the Agisoft PhotoScan framework. In a previous study, the authors showed that, in a trade-off between computation time and model quality, models can be generated using a single notebook within rather short time frames (less than 30 minutes) [1], allowing the process of drone programming, image acquisition and model processing to be conducted *in situ* in case of large-scale disaster events. In addition to the aerial images, the base of the water tower was

recorded terrestrially on a circular path around it through 57 images with a resolution of 6,000 x 4,000 pixels to demonstrate the depiction of details. Although conducted manually, the images could have been made using a drone at low level flight as well. The data set was processed with the proposed parameters and frameworks. The processing time of the open-source framework was about 12 hours with a necessary storage space of 25GB, with the Agisoft PhotoScan framework only about 4 hours with 10 GB of storage space. Both models are shown in Figure 8. Finally, high detail models like a shoe illustrated smaller objects (Fig. 9) can be included to the model using photogrammetric imaging utilizing motor-driven turntables.

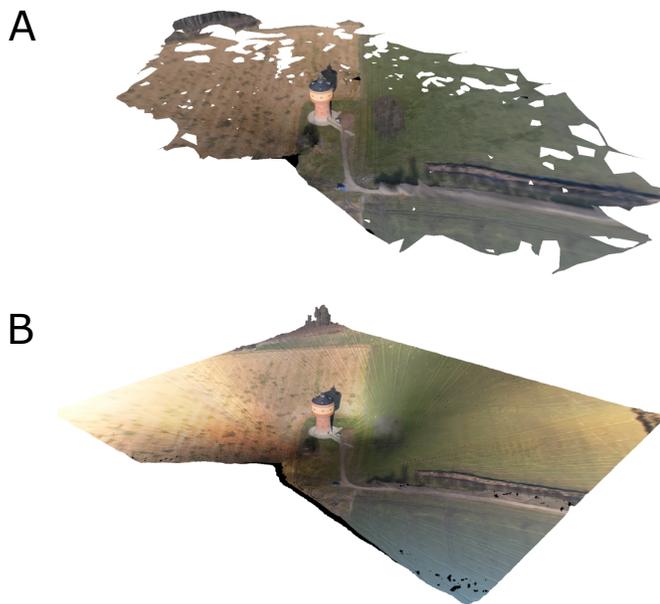


Figure 7. Watertower of Mittweida based on 111 UAV-images processed using the open source (A) and Agisoft PhotoScan (B) software.

C. Test Scenario II — Crime Scene Reconstruction in Jena Lobeda

The crime scene in question addressed in this scenario is located in a remote area of Jena Lobeda city directly adjacent to the Saale river. 3D reconstruction of the location showed to be demanding due to the site being difficult to access. In addition multiple obstacles made optimal drone flight geometry impossible to achieve. Thus, the site had to be inspected beforehand and careful planning had to be carried out. For example obstacle heights were determined during the inspection and confirmed from estimations based on areal images. Consistent visual contact to the drone had to be ensured as well (as specified by German law). In Figure 10 a schematic of the area is shown (subfigure A) and a detail with annotated minimum and maximum safe altitudes is shown. The location of obstacles as well as the remoteness of the crime

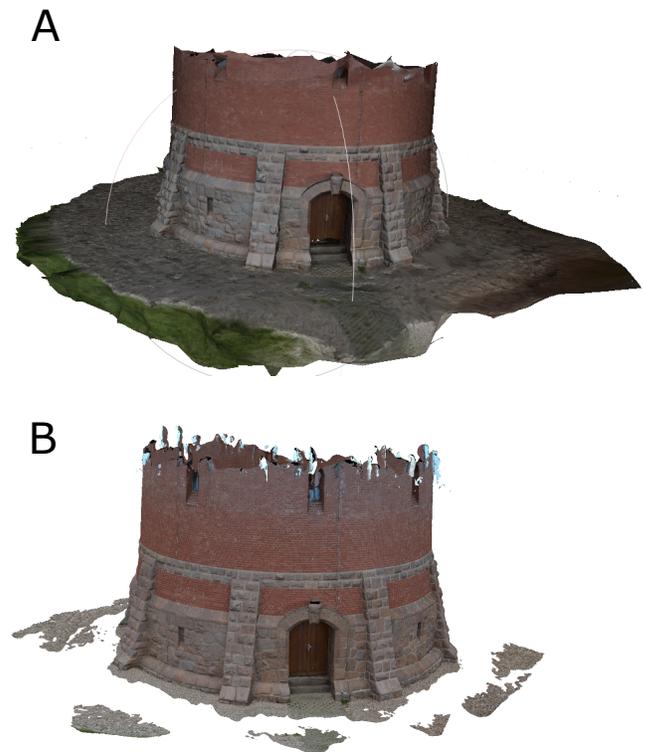


Figure 8. Watertower of Mittweida based on 57 terrestrial images processed using open source (A) and Agisoft PhotoScan (B) software.

scene only allowed the drone to be setup as well as to take off and land on the western riverbank. Furthermore, a flight plan was chosen, which provided imagery of the area of interest by flying parallel to the eastern riverbank. Multiple flight legs were conducted during this flight phase, each with individual preprogrammed camera angles and flight altitudes. The final flight phase included a 45° look-down camera angle to the crime scene from the other side of the area. In Figure 10B the chosen flight path is depicted schematically. A screenshot of the MikroKopter Tool showing the actual flight path is given in Figure 10C. During flight 120 single images were extracted from HD video recordings taken with a drone-mounted SLR camera. In addition a Eurocopter EC-145 helicopter equipped with a StarSAFIRE HD camera (resolution: 1,280 x 720 pixels) was provided by the Jena police department. 250 images were extracted from the video file recording of the helicopter-mounted camera. Laser scanning of the crime scene was conducted as well, but no satisfying data could be gathered due to dense and tall-grown vegetation and laser scattering effects caused by prevalent humidity. Hence laser scanning showed to be inferior to photogrammetric techniques under the present conditions.

D. Software Utilization and Parameterization

1) *VSFM-CMPMVS-MeshLab*: Prior to importing the images to VisualSFM, it is important ensure that the images



Figure 9. Digitized shoe (based on 30 terrestrial images) as example for a seized trace at a crime scene.

are of proper quality. Especially in case of extracting images from video recordings, the imagery can be prone to containing unwanted artifacts, such as motion blur or lense flares. In this respect, the VisualSFM GUI provides a convenient presentation interface of all images loaded to the workspace from, which images lacking proper quality can be removed. With all images being selected, parameterizations of point cloud calculation is conducted. It should be highlighted that the primary parameter in calculation is the maximum number of underlying SIFT features to be considered. The point cloud quality can thus be increased by raising the number of SIFT features, whereas time and memory demands however increase significantly. Another aspect to address is that images with resolution greater than 3,200 x 3,200 pixels are down-scaled by default, decreasing point cloud quality. In all presented scenarios this setting was changed accordingly in order to avoid downscaling. Again these selected settings are accompanied with significant computational demands. Upon SIFT feature and point cloud calculation, storing the data in the appropriate file format (option "CMPMVS") provides a CMPMVS-compatible initialization file, which in turn can be feed to the CMPMVS executable. After computation the resulting model is stored in a WRL file for further processing. As the next step, the authors recommend to employ MeshLab for inspecting the obtained model (Fig. 11, subfigure A and Fig. 12, subfigure A) with respect to general quality and artifacts. MeshLab can also be employed to perform first model corrections and removal of existing artifacts.

2) *Agisoft PhotoScan*: The framework (Fig. 6) for creating 3D scenarios using the commercial software application Agisoft PhotoScan (Professional Edition, version used in this study: V1.4) can be subdivided into the two consecutive

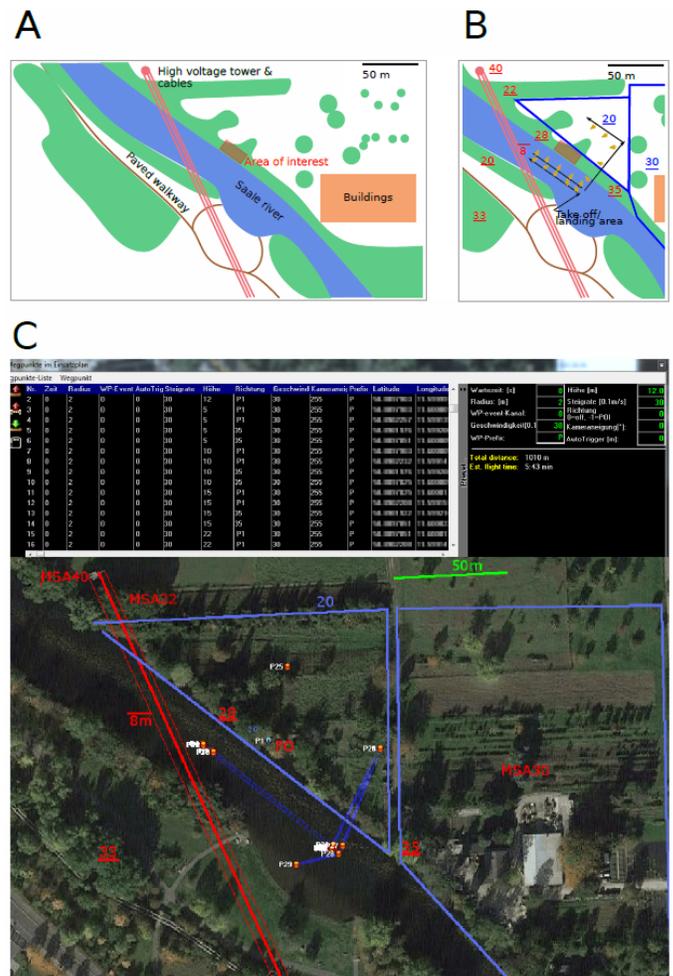


Figure 10. Steps involved in drone flight planning for the Lobeda crime scene reconstruction. A: Schematic overview of the area in question. Multiple obstacles of varying height (f.e. trees and a high voltage power line) hinder the application of a circular flight path. The crime scene area had thus to be inspected beforehand in this case. In (B) height information gathered from inspection and areal images is annotated (red: maximum and minimum safe obstacle heights, blue: minimum safe height areas) and the flight is planned accordingly. Black and yellow arrows indicate the drone flight path and camera angles respectively. C: Screenshot of the MikroKopter Tool showing the loaded flight path.

modules "Pre-Processing" and "Model Generation", which in turn consists of specific sub-processes. The initial phase of the "Pre-Processing" module consists of the selection and import of the previously obtained images. As with the utilization of VisualSFM, the user has to ensure proper file formatting and quality. In its current version Agisoft PhotoScan also supports the direct processing of video data. Here convenient video import and eventual image extraction is available. Similar to VisualSFM images are displayed in a dedicated graphical interface for quality assessment and selection. The information about the resolution of the images and/or EXIF data are essential to decide whether a camera calibration is necessary. Camera positions are computed next. However, considering

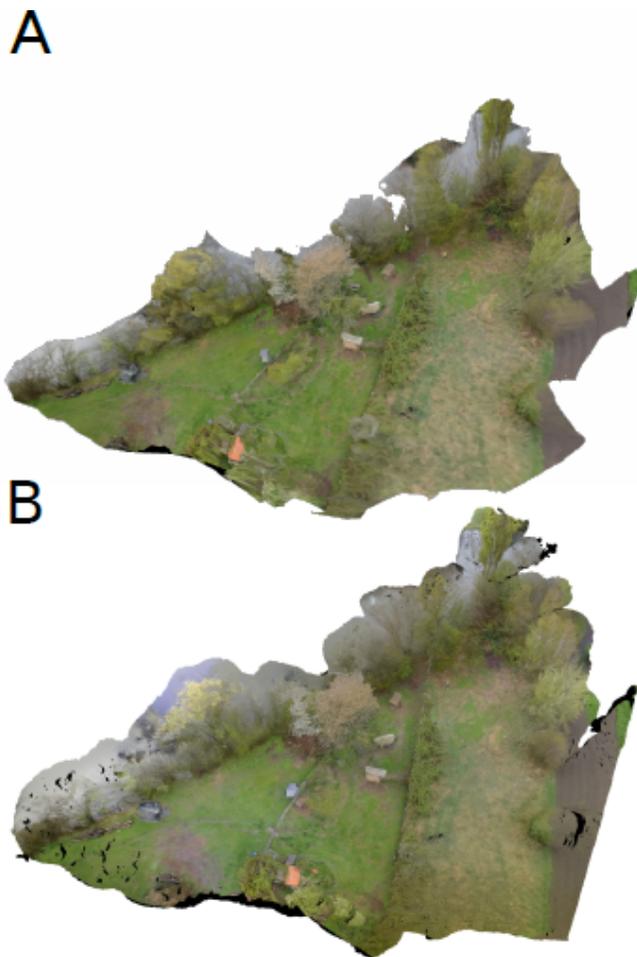


Figure 11. Saale in Jena based on UAV-images processed using the open source (A) and Agisoft PhotoScan (B) framework.

the data collection conditions and parameterizations presented in the “Data Collection” framework, this processed (referred to as a manual camera calibration) can be avoided in most cases. In the forensic context it is important that the created model has a high quality. This can be achieved in Agisoft PhotoScan by using the maximum values for all parameters. However, the resulting computation afford makes it necessary to decrease some of the values while keeping the quality of the model high. In own experiments sets of parameters (see Table I) have been evaluated and used in creating models with a sufficient quality for subsequent simulations. After creating the 3D models, they were edited and textured (Fig. 11, subfigure B and Fig. 12, subfigure B).

E. Results

It was shown that helicopters are only partly suitable for the aerial photogrammetry (Fig. 12). Possible scenarios in, which a helicopter can be of advantage are reconstructions of huge areas without high level of degree. Disadvantages in the

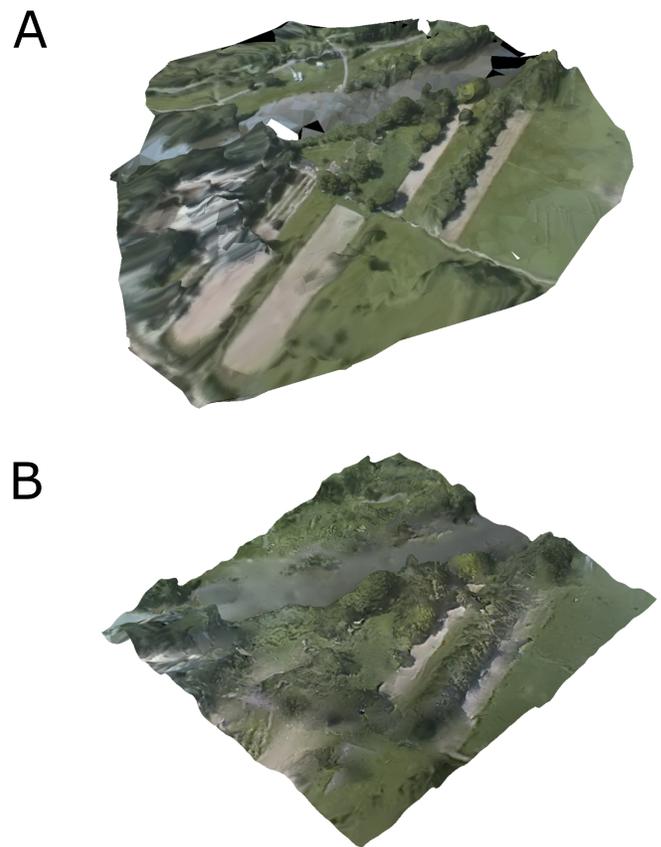


Figure 12. Saale in Jena based on helicopter images processed using the open source (A) and Agisoft PhotoScan (B) framework.

TABLE I. Recommended specific parameter sets for the generation of models in Agisoft PhotoScan.

task	parameter	min	max	recommended
aligning	tie point limit	0	n	10,000
	key point limit	0	n	400,000
dense cloud mesh	accuracy	lowest	highest	high
	surface type			Arbitrary
	face count	low	custom	default

described scenarios were next to the low resolution of the camera, especially the relatively high flight altitude. Because the drone can fly in lower altitudes and the payload can be chosen flexibly a high level of degree can be ensured. Additionally, the compact design allows its operation in many scenarios. The only limiting factor besides the weather is its limited battery service life. Doing a reconstruction using terrestrial photographs is suitable for the reconstruction of objects with an extremely high level of degree and a manageable size. However, difficult terrain and obstacles (e.g., branches or high grasses) can lead to problems.

IV. CONCLUSION AND FUTURE WORK

The 3D reconstruction of objects, complex crime scenes or disaster sites provide a quick and comprehensive method to support the investigators digitally. Photogrammetry-based methods to creating 3D models are an inexpensive alternative to laser scanners. Even though various approaches exist, to this point there is no standardized process for the reconstruction of crime scenes or disaster sites. Therefore, in this paper the authors present a comprehensive framework for the 3D reconstruction process using terrestrial and aerial photogrammetry. The advantage of the presented framework is twofold. It can serve as a simple guide to create a 3D reconstruction, yet it can also be easily implemented in an already existing process chain, for example as used by investigative services. Additionally, the complete framework from the planning phase up until the final 3D model is presented by means of a test scenario as well as one cold case. It could be shown by the authors, that 3D models of sufficient quality for subsequent simulations can be generated using the presented framework. With respect to the two presented frameworks we focused on discussing model quality and processing parameterization. In practical use one has to consider these aspects as a trade-off of computational demands and resulting model resolution and quality. Future work thus has to address the definition and evaluation of quality measures as well as eventual quantification of best-practice standards, including guidelines for optimizing modelling strategies. Future prospects should also consider a quantitative comparison to established laser-scanning strategies. Finally the adaption to a wide range of practical usage scenarios by means of utilizing alternative software packages and/or payloads (e.g., thermal imaging or gas sensors) could be discussed.

ACKNOWLEDGMENT

The authors would like to thank the Free State of Saxony for funding and the Jena police department for providing the opportunity to work on a real case.

REFERENCES

- [1] M. Spranger, F. Heinke, S. Becker, and D. Labudde, "Towards drone-assisted large-scale disaster response and recovery," ACCSE, 2016, pp. 5–10.
- [2] A. Marcin, S. Maciej, S. Robert, and W. Adam, "Hierarchical, three-dimensional measurement system for crime scene scanning," *Journal of forensic sciences*, vol. 62, no. 4, 2017, pp. 889–899.
- [3] E. Hołowko, K. Januszkiewicz, P. Bolewicki, R. Sitnik, and J. Michoński, "Application of multi-resolution 3d techniques in crime scene documentation with bloodstain pattern analysis," *Forensic science international*, vol. 267, 2016, pp. 218–227.
- [4] D. Dustin, E. Liscio, and P. Eng, "Accuracy and repeatability of the laser scanner and total station for crime and accident scene documentation," *J Assoc Crime Scene Reconstr*, vol. 20, 2016, pp. 57–68.
- [5] U. Buck, S. Naether, B. Räss, C. Jackowski, and M. J. Thali, "Accident or homicide—virtual crime scene reconstruction using 3d methods," *Forensic science international*, vol. 225, no. 1–3, 2013, pp. 75–84.
- [6] V. Pagounis, M. Tsakiri, S. Palaskas, B. Biza, and E. Zaloumi, "3d laser scanning for road safety and accident reconstruction," in *Proceedings of the XXIIIth international FIG congress*, 2006, pp. 8–13.
- [7] U. Buck, "Laserscanning in der kriminalistik," *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, vol. 135, 2010, pp. 190–198.
- [8] H. Püschel, M. Sauerbier, H. Eisenbeiss et al., "A 3d model of castle landenberg (ch) from combined photogrammetric processing of terrestrial and uav-based images," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 37, 2008, pp. 93–98.
- [9] M. Douglass, S. Lin, and M. Chodoronek, "The application of 3d photogrammetry for in-field documentation of archaeological features," *Advances in Archaeological Practice*, vol. 3, no. 2, 2015, pp. 136–152.
- [10] K. Porter, R. Simons, and J. Harris, "Comparison of three techniques for scour depth measurement: photogrammetry, echosounder profiling and a calibrated pile," *Coastal Engineering Proceedings*, vol. 1, 01 2014, p. 64.
- [11] A. Morales, D. Gonzalez-Aguilera, M. A. Gutiérrez, and A. I. López, "Energy analysis of road accidents based on close-range photogrammetry," *Remote Sensing*, vol. 7, no. 11, 2015, pp. 15 161–15 178.
- [12] A. Morales, L. J. Sánchez-Aparicio, D. González-Aguilera, P. Rodríguez-González, D. Hernández-López, M. A. Gutiérrez, and A. I. López, "A new approach to energy calculation of road accidents against fixed small section elements based on close-range photogrammetry," *Remote Sensing*, vol. 9, no. 12, 2017, p. 1219.
- [13] P. Urbanova, P. Hejna, and M. Jurda, "Testing photogrammetry-based techniques for three-dimensional surface documentation in forensic pathology," *Forensic science international*, vol. 250, 2015, pp. 77–86.
- [14] T. P. Kersten, C. A. Pardo, and M. Lindstaedt, "3d acquisition, modelling and visualization of north german castles by digital architectural photogrammetry," in *Proc. ISPRS XXth Congress*, 2004, pp. 126–131.
- [15] T. P. Kersten and M. Lindstaedt, "Generierung von 3D-Punktwolken durch kamera-basierte low-cost Systeme ? Workflow und praktische Beispiele." *Terrestrisches Laserscanning 2012*, vol. 69, 2012, pp. 25–46.
- [16] —, "Image-based low-cost systems for automatic 3d recording and modelling of archaeological finds and objects," in *Euro-Mediterranean Conference*. Springer, 2012, pp. 1–10.
- [17] P. L. Falkingham, "Acquisition of high resolution three-dimensional models using free, open-source, photogrammetric software," *Palaeontologia electronica*, vol. 15, no. 1, 2012, p. 15.
- [18] M. Ziegler, E. Gülch, and P. Rawiel, "3D-Rekonstruktion von Objekten mittels Structure-from-Motion aus einer photogrammetrischen Aufnahme mit den Programmen VisualSFM und CPMVMS," in *DGPF Tagungsband*, vol. 23, 2014.
- [19] I. Colomina and P. Molina, "Unmanned aerial systems for photogrammetry and remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 92, 2014, pp. 79–97.
- [20] M. R. Resig, "Rapid 3D Scene Reconstruction from Kite-Based Aerial Imagery Using Open Source Structure from Motion," Master Thesis, George Mason University, 2015.
- [21] H. Bendea, P. Boccoardo, S. Dequal, F. Giulio Tonolo, D. Marenchino, and M. Piras, "Low cost uav for post-disaster assessment," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, no. B8, 2008, pp. 1373–1379.
- [22] Y. Naidoo, R. Stopforth, and G. Bright, "Development of an uav for search & rescue applications," in *AFRICON, 2011*. IEEE, 2011, pp. 1–6.
- [23] L. Barazzetti, R. Sala, M. Scaioni, C. Cattaneo, D. Gibelli, A. Giussani, P. Poppa, F. Roncoroni, and A. Vandone, "3d scanning and imaging for quick documentation of crime and accident scenes," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI*, vol. 8359. International Society for Optics and Photonics, 2012, p. 835910.
- [24] S. Zancajo-Blazquez, D. Gonzalez-Aguilera, H. Gonzalez-Jorge, and D. Hernandez-Lopez, "An automatic image-based modelling method applied to forensic infography," *PLoS one*, vol. 10, no. 3, 2015, p. e0118719.
- [25] C. Kim, H. Moon, and W. Lee, "Data management framework

- of drone-based 3d model reconstruction of disaster site,” ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLI-B4, 2016, pp. 31–33. [Online]. Available: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B4/31/2016/>
- [26] D. Gonzalez-Aguilera, L. López-Fernández, P. Rodríguez-Gonzálvez, D. Hernández-López, D. Guerrero, F. Remondino, F. Menna, E. Nocerino, I. Toschi, A. Ballabeni et al., “Graphos—open-source software for photogrammetric applications,” *The Photogrammetric Record*, vol. 33, no. 161, 2018, pp. 11–29.
- [27] D. Gonzalez-Aguilera and J. Gomez-Lahoz, “Forensic terrestrial photogrammetry from a single image,” *Journal of forensic sciences*, vol. 54, no. 6, 2009, pp. 1376–1387.
- [28] K. Makantasis, A. Doulamis, N. Doulamis, and M. Ioannides, “In the wild image retrieval and clustering for 3d cultural heritage landmarks reconstruction,” *Multimedia Tools and Applications*, vol. 75, no. 7, 2016, pp. 3593–3629.
- [29] Z. Xu, T. Wu, Y. Shen, and L. Wu, “Three dimensional reconstruction of large cultural heritage objects based on uav video and tls data,” *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 41, 2016, p. 985.
- [30] R. K. Napolitano, G. Scherer, and B. Glisic, “Virtual tours and informational modeling for conservation of cultural heritage sites,” *Journal of Cultural Heritage*, 2017.
- [31] J. A. Torres-Martínez, M. Seddaiu, P. Rodríguez-Gonzálvez, D. Hernández-López, and D. González-Aguilera, “A multi-data source and multi-sensor approach for the 3d reconstruction and web visualization of a complex archaeological site: The case study of “tolmo de minateda”,” *Remote Sensing*, vol. 8, no. 7, 2016, p. 550.
- [32] C. Wu. SiftGPU Implementation of Scale Invariant Feature Transform. [Online]. Available: <http://cs.unc.edu/ccwu/siftgpu/>
- [33] C. Castillo, M. James, M. Redel-Macías, R. Pérez, and J. Gómez, “Sf3m software: 3-d photo-reconstruction for non-expert users and its application to a gully network,” *Soil*, vol. 1, no. 2, 2015, p. 583.
- [34] S. Becker, J. Dreßler, K. Thiele, and D. Labudde, “Gesichtswerteilrekonstruktion mithilfe einer open-source-software,” *Rechtsmedizin*, vol. 26, no. 2, 2016, pp. 83–89.
- [35] M. Jancosek and T. Pajdla, “Cmpmvs-multi-view reconstruction software,” in *Proceedings of the International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, 2008.
- [36] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, “Meshlab: an open-source mesh processing tool.” in *Eurographics Italian Chapter Conference*, vol. 2008, 2008, pp. 129–136.

MobileTimeSync - An Android App for Time Synchronization for Mobile Construction Assessment

Maik Benndorf, Marika Kaden,
Frederic Ringsleben and Thomas Haenselmann
Faculty of Applied Computer
Sciences & Biosciences
University of Applied Sciences
Germany, Mittweida 09648
{benndorf, kaden1, ringsleb, haenselm}@hsmw.de

Eric Zuchantke
Faculty of Mathematics and
Computer Science
Friedrich-Schiller-Universität Jena
Germany, Jena 07743
eric.zuchantke@uni-jena.de

Martin Gaedke
Faculty of Computer Science
Technische Universität Chemnitz
Germany, Chemnitz 09111
gaedke@informatik.tu-chemnitz.de

Abstract—Modern smartphones offer more use-cases than (just) pure communication. These devices, equipped with several sensors, can be used as measuring devices, for example. In order to compare and interpret the sensor data obtained from multiple devices, precisely synchronized clocks are needed. Since all clocks are subject to inaccuracies, a synchronization method is necessary. In this article, we propose MobileTimeSync - an Android app for time synchronization. We suggest how to align clocks based on ultrasound beacons to mutually take clock drift and skew into account. In addition, we will discuss the different kinds of time sources in Android smartphones. With our approach, we obtain a precision, which can be expected from traditional synchronization approaches such as Network Time Protocol but with much less effort.

Index Terms—time synchronization; clock skew; clock drift; sensor networks; mobile devices; smartphones

I. INTRODUCTION

Most computer clocks, e.g., those built into a smartphone, are subject to an inaccuracy [1]. This inaccuracy can be neglected in the general purpose of the smartphone. However, since these devices are equipped with several sensors, there are other use cases in which this inaccuracy cannot be neglected.

The following scenario illustrates our need for accurate clocks: In cases of natural disasters such as flooding, timely access to information about transport infrastructure is crucial for the first-aiders as well as for the affected people. In particular, bridges could be fit for traffic, partially usable or be destroyed entirely. While in many cases a visual inspection is sufficient to assess the accessibility of roads, for the assessment of bridges further information is needed. Assessing the construction's integrity can be accomplished by the analysis of the bridge's eigenfrequencies. Therefore, a proprietary, expensive Vibration Measurement Systems (VMS) is necessary. However, in many cases, such a device is not available in a timely manner in the areas struck by a disaster. In [2], [3], we were able to show how this VMS can be replaced by low-cost, off-the-shelf smartphones that can be applied to perform an ad-hoc assessment of a bridge. Our current research is now focusing on the issue of locating the damage on the bridge. For this purpose, several smartphones have to be placed on

the surface of the bridge, and their measured values have to be aligned finally. Therefore, the clocks of these devices must be synchronized.

At the Sensorcomm conference we presented an approach for acoustic time synchronization of smartphones [1]. The basic idea for this synchronization approach comes from sports. With a 100 m sprint, eight athletes stand in starting blocks. The race begins with an external signal which is the same for each athlete. This idea was adapted for the time synchronization of smartphones. For that purpose, an external acoustic signal is generated and sent out. The smartphones microphone receives this signal and the timestamp of the arriving of the signal is used for aligning different devices. If several of those signals were sent out with this procedure, it is possible to determine the clock skew and clock drift.

One issue of our previous work was the source of the sound signal. In our experiments in [1], we used a choke in the same way as it is applied in athletic sports. In our recent research, we create an Android app (see Fig. 1) for this purpose. This app serves as both a sender (sound source) and a receiver. Moreover, the app facilitates the use of time synchronization on these devices and will be presented in this paper.

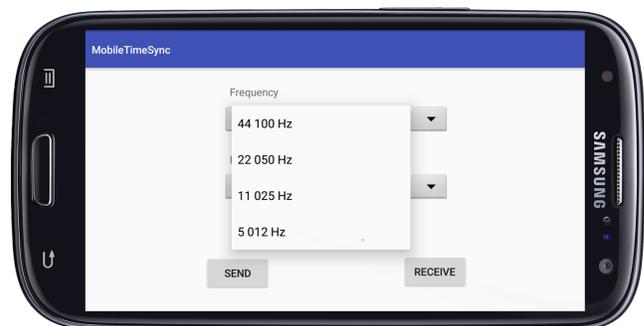


Fig. 1. Screenshot of the Android App for Time Synchronization

Therefore, this article is organized as follows: Section II is devoted to provide fundamentals and give an overview of related work in the field of synchronizing methods. Our application and methodology is described in Section *Methodology*

(Section III). Some experiments and their results are presented in Section IV. A summary and outlook are drawn in Section V.

II. FUNDAMENTALS AND RELATED WORK

The addressed problem is related to the time synchronization in sensor networks. A lot of research has been done in this area. There are several methods for synchronizing physical clocks. These methods can be classified as "internal" and "external" synchronization. For an internal clock synchronization, all nodes accept the time of a reference node in the network [4]. In external clock synchronization, the value is taken from an external clock source, such as a common time service, e.g., Network Time Protocol NTP or the Global Positioning System (GPS) [5].

A. Synchronization Methods

In this section, some of the most common synchronization methods are shortly discussed.

1) *Synchronization by NTP*: NTP is a protocol for synchronizing computer clocks using a set of distributed servers around the world. This protocol is also known as Simple Network Time Protocol (SNTP). It is built on top of the User Datagram Protocol (UDP) [6]. Fig. 2 shows the process of an NTP based time synchronization. The protocol was announced with a precision in the range of nanoseconds (ns) [7] [8]. This protocol has been utilized in numerous clients for several years. Juda Levine [9] reports in 2011 about 5×10^9 requests per day. The accuracy of the protocol and the related assets have been studied in numerous works [10]–[12]. The network latency has a significant impact on the accuracy. Zhao et al. [11] evaluated the accuracy with less than 10 ms under Local Area Network (LAN) condition and less than 100 ms under Internet conditions. As a reference for their evaluations, they used the time of the GPS.

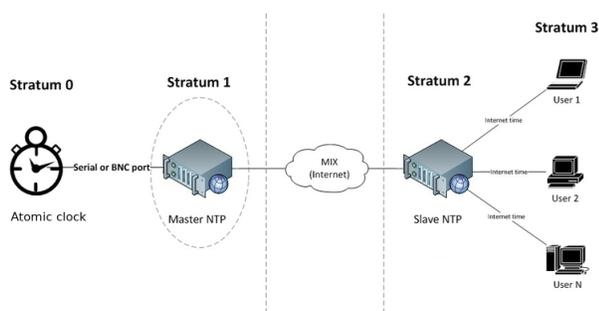


Fig. 2. Overview of time synchronization via NTP [13]

2) *Synchronization by GPS*: The Global Positioning system, was designed and is still under the control of the United States Department of Defense. Nevertheless, it is also freely accessible by anyone. The system consists currently of 32 operating satellites that are orbiting the earth at an altitude of approximately 20,000 km. Every satellite contains multiple atomic clocks that support very precise timing data [14]. For

determining the position, the receiver needs signals from at least three of these satellites. The position of the receiver can be calculated by the difference between the signal sent and received by the receiver. With this ability to receive very accurate data from multiple sources an accurate time can be obtained. In 2015, Mazur et al. [14] determined the accuracy of such time synchronizations within up to billionth of a second. This system can be accessed anywhere in the world and has very high accuracy. But it requires a direct line of sight to the satellite, and an initial connection takes a long time in some cases.

3) *Time Synchronization on Mobile Devices*: One recent work in this area is provided by Sridhar et al. In [15] they proposed Cheepsync - a synchronization service for Bluetooth Low Energy (BLE) Broadcasters. Cheepsync consists of a custom transmitter platform and an Android smartphone as the receiver. In order to overcome the non-deterministic delays, they used a common BLE chip. Since the transmission command directly addresses the hardware, they were able to generate a low-level timestamp in the range of 10 μ s.

Another approach is proposed by Lazik et al. [16]. They used ultrasonic beacons to synchronize the time on mobile devices. Therefore, they built up a network with one network master. The master is connected to a GPS receiver and transmits ultrasonic chirps in a frequency that is outside of the human hearing but still detectable by the microphone of smartphones. They reported that the devices could be synchronized with an average accuracy of 720 μ s. At the beginning of their experiments, they investigated devices with Android and iOS. They reported a high-level jitter on the Android device (in the order of milliseconds and higher) and chose the iOS devices for the rest of their experiments. The jittering is justified by Android's task scheduler. Within this work, they also benchmarked the NTP timing performance (on iOS). They ran their experiments with three different communication channels: Long Term Evolution Technologie (LTE), Wireless Local Area Network (WLAN), and one idle WLAN router that is directly connected to a Stratum 1 NTP server fed by a dedicated GPS clock. Using LTE they measured an average jitter of 47 ms, the average WLAN jitter is measured with 30 ms, and finally, the average jitter in the ideal case with the WLAN router connected to NTP Server is measured with 19.3 ms.

B. Available Time Sources on Android Devices

In Android devices different time sources are provided. A distinction can be made between the hardware timers which exist in every device, the timers offered by the Operating System (OS) and finally the timers provided by Java Virtual Machine. To use the correct time source for the synchronization, these sources are explained in more detail in the following.

1) *The Hardware Timers*: Accessing hardware timers on Android devices can be useful if the application only has to run on targeted hardware. It can be rewarding to do so

because these timers are incredibly accurate, monotonic, non-decreasing clocks and cheap to read. On some ARM CPUs the overhead to read the time can be as small as only one CPU cycle (e.g., for Cortex A9 processors). But this high performance comes at a price as directly accessing the hardware requires an assembly code which is not portable, takes a lot of time to develop and is hard to read. Furthermore, if the program ought to be used on different devices most of them will have different hardware clocks. Errors can also occur when the device enters CPU saving or low-frequency states. All in all this method would only be suitable if the hardware is precisely targeted, which is not in our case [17].

2) *Android Based Timers:* Times from the Android operating system can be obtained by using clock sources from the Linux kernel. As in any system based on the 2.6 kernel, Android holds a linked list of all its clock sources sorted by rating and chooses the best available clock for all functions getting time from the kernel. One function to access these time sources is the `clock_gettime()`-posix function, which returns the time in a resolution of ns. However, this produces a small, acceptable overhead on ARM CPUs. If the `clock_gettime()`-function is running on a system which implements a syscall for it the overhead can also be removed. But as long as this method is only usable in native applications developed with C or with the usage of the Java Native Interface (JNI), it may not be suitable for particular use cases [17].

3) *Java Timers:* Because many Android applications are written in Java, the Java time functions are also considered for use as they are implemented in the Java Virtual Machine (JVM) and therefore are easily accessible from any Java application. There are four different types of clock functions offered that may prove useful [18]:

- `SystemClock.uptimeMillis()`
- `System.currentTimeMillis()`
- `System.currentThreadTimeMillis()`
- `SystemClock.elapsedRealtime()`

`uptimeMillis()` returns the system uptime since last boot in milliseconds (ms). The clock is guaranteed to be monotonic and is unaffected by most power saving mechanism, but it stops when the device enters deep sleep [18]. Therefore, this can be a reliable time source only if the measured time does not span deep sleep. However, due to the possibility of device going into 'deep sleep mode' this clock is not the best choice.

`currentTimeMillis()` returns the time in ms since the epoch. It can be used to measure daytime and dates but can be set by the user or the phone network. So the returned time may jump around [18]. In conclusion, the time provided by this function would be great to use in the application, but is too unreliable due to time changes.

`currentThreadTimeMillis()` returns milliseconds running in the current thread [18]. However, this can only be

used for time measurement in combination with timestamps received from other time sources. Even if it is used that way, this will produce more overhead for time calculation than using `uptimeMillis()`. So, it is not considered as a helpful time source to compare timestamps among devices.

The functions `elapsedRealtime()` and `elapsedRealtimeNanos()` return the time since the system's last boot in ms and ns. Both functions include deep sleep and are guaranteed to be monotonic not to be influenced by the CPU entering power saving modes [18]. Thus, it can be useful for time comparison, but it should be noted that this is a relative time and cannot be used for comparison across different devices.

4) *Advanced Time Calculation with Combined Methods:* Another method for achieving accurate timestamps is called TrueTime [19]. The proposed algorithm takes advantage of the function `elapsedRealtime()` by requesting a timestamp from an NTP Server and storing this timestamp along with the `elapsedRealtime()` upon receiving the NTP timestamp in the device memory. Now every time a timestamp is needed the stored `elapsedRealtime()` can be subtracted from the current `elapsedRealtime()` and this difference is then added to the stored NTP timestamps creating the new current time in ms. Advantages of this method are that these timestamps are impervious to device clock changes as the NTP server dictates the time and the accuracy of this procedure above multiple devices is very high as long as every device receives the same timestamp via NTP. Also, the created times will be accurate until the device reboots, and it does not require a stable data connection. Disadvantages include the need for a data connection every time after reboot before the method can be used and the precision of the timestamp received that may vary depending on the latency of the network connection.

5) *Summary:* In summary, there are two types of times we can get from the system. First of all, there is the daytime which is dependent on which time the device was manually set to by the user or automatically.

Also, there are methods that will return timestamps measured relative to certain events like system boot. These methods deliver very accurate timestamps, but they are not comparable across different devices. But this is not necessary for our synchronization method. To compare these methods, we conducted an experiment with one of our phones. Therefore, we compared the Functions `System.currentTimeMillis()` and `SystemClock.elapsedRealtimeNanos()`. The mean difference between both the time sources is 0.1906 ms (± 0.3919). The mean difference between ground truth and `elapsedRealtimeNanos()` is 9.020 ns (± 7.779). And finally, the mean difference between ground truth and `currentTimeMillis()` is 9.088 ms (± 7.79). The results can also be seen in Fig. 3. The function `elapsedRealtimeNanos()`

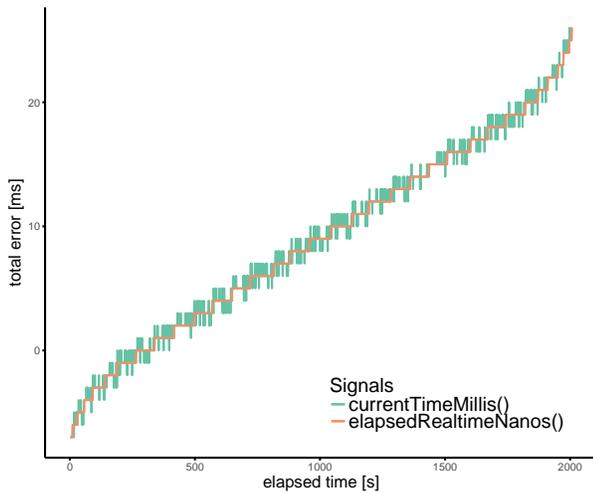


Fig. 3. Comparison of the two signals: (sorted) difference of ground truth to elapsedRealtimeNanos() (orange) and currentTimeMillis()(green)

performs only marginally better than the function currentTimeMillis(). Nevertheless, since only the function elapsedRealtimeNanos() is monotonic and not influenced by deep sleep, we decide to use this time source.

C. Problems

In combination with the time synchronization on Android devices, there are some problems which are mentioned below.

1) *Clock-Skew and Drift*: This is not an Android specific problem but rather a general problem of most of the computing devices. These devices are equipped with a hardware oscillator assisted computer clock. The frequency of the hardware oscillator determines the rate at which the clock runs [20]. This clock becomes inaccurate because the frequency varies.

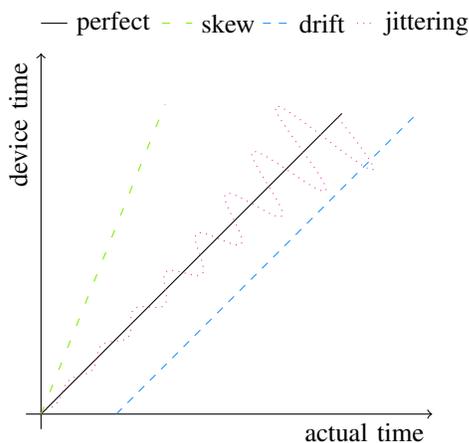


Fig. 4. The difference between a perfect clock, one with drift, one with a skew and one, as it is likely to occur in real [21].

Fig. 4 shows the difference between the clock drift - in this case, the clock is behind or ahead by a fixed offset, the clock

skew - here the offset is growing with the time, and the jittering - in this case, the device clock is affected by internal (e.g., processor utilization) or external (e.g., temperature, humidity [22]) fluctuations. To show the appearance of the clock skew with an example, we present here a result of an experiment shown in [1]. In this experiment, we compared the local time on three different smartphones over 24 hours with the time values of the GPS sensor.

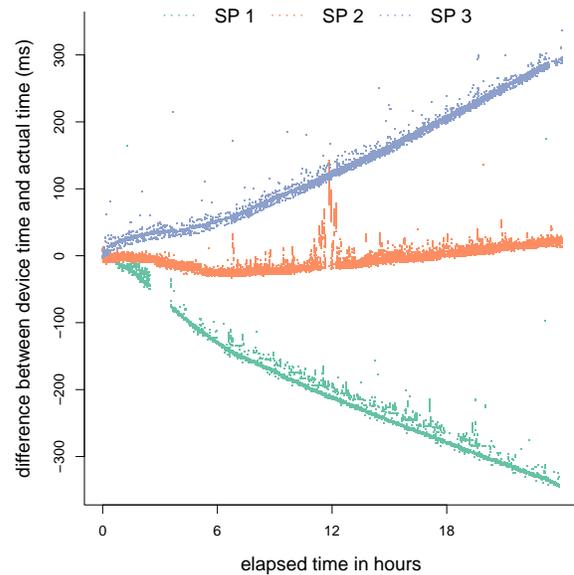


Fig. 5. The clocks skew of the devices within 24 hours compared to the timestamp given by GPS.

As it can be seen in Fig. 5 the offset of all three devices is growing over time. To keep these clocks in time, Zhenjiang Li et al. [23] use the flickering lights of fluorescent lights.

2) *Jittering*: Jittering is an Android specific problem. The duration of the execution of a function can vary. If the execution is heavily dependent on time measurement (like the comparison of sensor values in short amounts of time), these timing errors can lead to all kinds of possible misfunctions of the algorithm. So, it is advisable to look into the problems which can be caused by timing errors. The difference in the function execution time is called Jitter. According to an analysis from the University of Florida, the Jitter increases when polling device sensors at higher frequencies [24]. Data can be received from the device sensors via the sensor manager, which also creates timestamps for the data. However, this can unfortunately be received from nonmonotonic clock (creating problems for comparison of sensor data) as some hardware manufacturers do not fulfill the hardware specifications for sensor polling from Google. Sources for Jitter on Android devices are issues with the Android Compatibility Definition as it lacks hardware standards and requirements like minimizing Jitter are only "should" criteria. Moreover, it can be constrained from the devices resources. These are mainly Memory (Garbage Collection / Thread Queues), Processor (Scheduler / Interrupt Handling) and Energy. To

reduce resulting issues, developers can force the sensors to synchronize with the `elapsedRealtimeNanos()`-clock from the Android SDK to get real-time timestamps for the sensor data as this is not guaranteed for every device. In case of data received from multiple sensors, the study recommends to create timestamps for every sensor individually and synchronize the data afterwards [24]. Monotonic time series can only be reliably achieved through manual synchronization with a strictly monotonic clock as the requirements for Android. This also gives the possibility to synchronize sensors with non-monotonic clocks. Also, event delivery delay can be reduced by ensuring that the device has enough available resources during the runtime of the application. However, this can be difficult to achieve as many consumer devices may suffer from resource starvation even without an application trying to reserve as many resources as possible. The ideal proposed alternative for developers is to find polling frequencies for the sensor at which Jittering cannot occur or stays within a forcible range [24]. To achieve these, developers can refer to the Android Compatibility Definition (CDD) for the respective android version.

III. METHODOLOGY

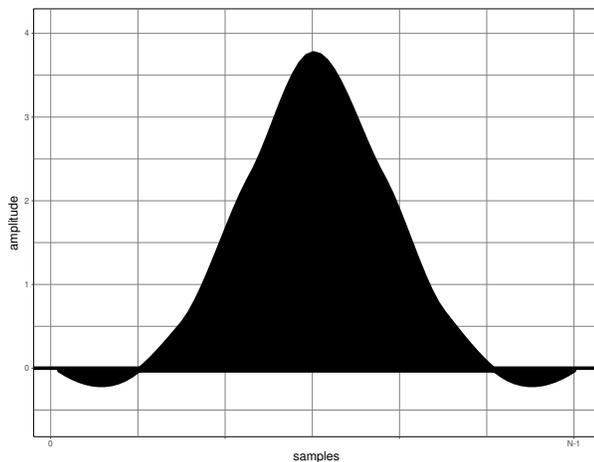


Fig. 6. Flat Window

In the previous section, we explained some synchronizing methods. Synchronization with a common time service will not be considered at the moment. This is because we can not make any assumptions about the availability, connection, and bandwidth. A prerequisite of synchronization among different devices is that they share the same network. To make the use case as simple as possible, it is not planned to create a network between those devices. Another option would be to synchronize via GPS. However, a direct view to the satellite must always be guaranteed. So, we decided to use acoustic synchronization. In [1] we created sound beacons with a choke from the sport. These acoustic beacons were recorded by the microphones of the smartphones. Afterwards, the peaks were isolated from the recorded data and the time at the peak was used to align the data of the various devices. In the

continuation of this work, we developed an app. This app (see Fig. 1) serves both as generator and transmitter of the peak and the receiver.

A. Specifics

Due to the operating system, there are some specifics that need to be mentioned.

- The sound is encoded by 16 bit. Hence, values in the range from $-32,768$ to $32,767$ with a maximum signal to noise ratio of 96.33 dB can be achieved.
- The audio samples are provided by the operating system as chunks. Therefore, only the time when a package is received can be measured. The time for the samples is calculated from the sample rate by interpolation. The size of the chunks depends on the buffersize that is proposed by Android.

In order to reduce external influences, we disable most of the applications on the device.

B. Sound Synthesis

To generate the synthetic peak, we use the Flat top window function. Fig. 6 shows one peak generated with this function. Depending on the selected sample rate and the selected periodicity, this peak is embedded in an audio file one or more times. The standard length of this audio file is about 30 minutes.

The synchronization process is started by pressing the send button. This will play the audio file and can be recorded by all nearby devices.

C. Sound Recording

All devices involved in synchronization must be close to the transmitter (to minimize the influence of sound). On these devices, the synchronization process is started by pressing the Receive button. In our example, we have chosen a sample rate of 44,100 Hz and a frequency of 1 Hz. This means that the peak was embedded in the audio file once per second. Fig. 7 shows (a) a single peak of the generated signal and (b) the sequence of several such peaks.

D. Processing

The following operations are performed in the background of the app.

1) *Interpolating the time for each sample:* As mentioned above, the data is only provided in chunks. This means that for each sample in this chunk, the times must be interpolated. One step for interpolation is calculated by the following equation:

$$I_s = \frac{(t_n - t_{n-1})}{S_r} \quad (1)$$

where I_s is the stepwidth for the interpolation, t_n is the timestamp the current chunk is received, t_{n-1} the timestamp the last chunk was received and S_r is the selected sample rate in our example 44,100 Hz.

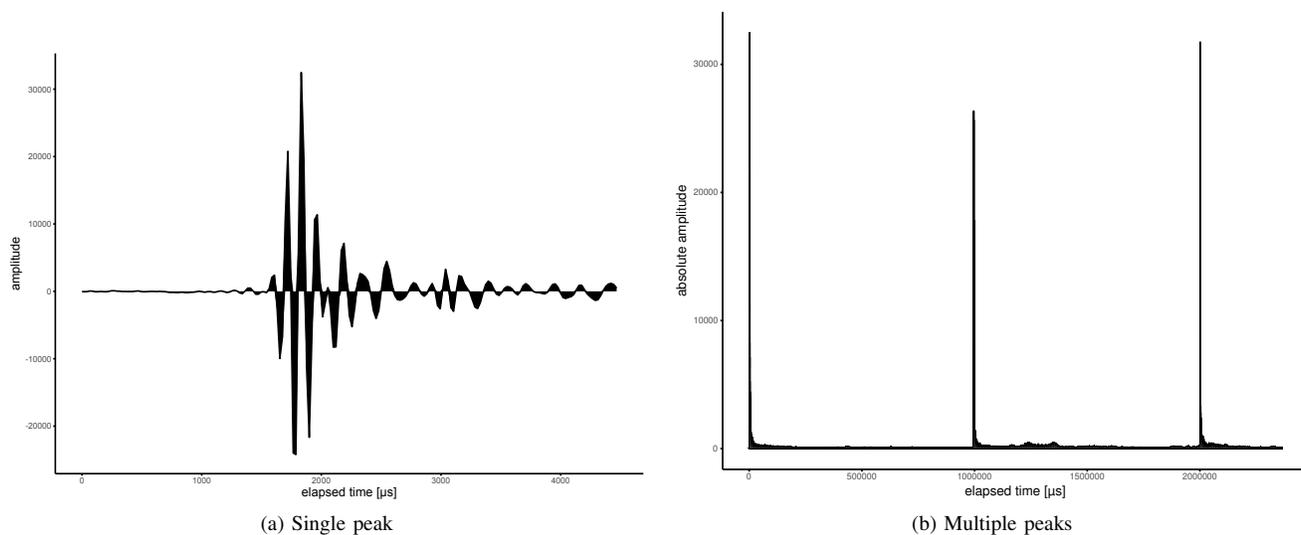


Fig. 7. App recorded (a) single peak or (b) multiple peaks.

2) *Preprocessing*: In the given signals of each device missing values can occur, i.e., the peak of the audio signal cannot be detected. In these cases the signal is replenished, i.e., if the difference between two records is greater than $periodicity \times 10\%$ the missing value is added in between. Thereby, the distance is fixed by $periodicity \times sample_rate$.

E. Filtering and storing the values

Afterwards, the values are filtered by amplitude. For this purpose, a threshold value can be configured in the App. The default value is 80 % of the maximum amplitude. All filtered values are stored in a database and can be used to align the devices.

IV. RESULTS AND EXPERIMENTS

During our experiments, we used different model devices from different manufacturers and will briefly denote this with Sp1, Sp2 and so on. These devices are equipped with up to three MEMS-Microphones (Micro-Electro-Mechanical-Systems) and are all operating with the Android operating system in versions from 4.2 up to 8.1.

A. Single Error Analysis for One Device

Fig. 8 shows, as an example for one of the devices, the simple error of the measured values during a half-hour synchronization process. The difference between ground truth (audio file) and the measured time of the received signal is colored gray and called single error. This deviation can be explained by the jittering of the operating system. The red line represents the mean value of the deviation (0.0105 ms (± 4.600)). Since the mean derivation is slightly positive, the deviations are accumulated, and thus, the total error increases over time.

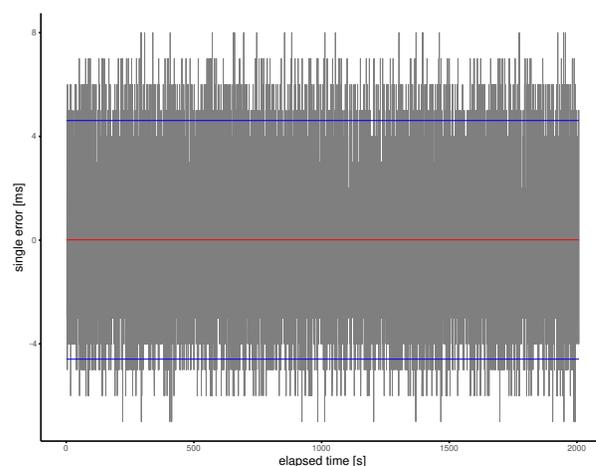


Fig. 8. Single derivation of the received signals to the ground truth of 1000 ms with the mean single error (red) and standard deviation (blue)

B. Single Error Comparison

All devices have similar behavior in the total error development (see Fig. 9). The deviation to the ground truth of Sp1 is less compared to especially Sp3. This can also be seen in the statistical values listed in Table I. Moreover, the single mean errors of each signal are small (about 0.01 ms). However, the standard deviations are high for all three devices (greater than 4 ms). This reflects the maximum of the single error, which is around 10 ms. These values correspond to those of [16].

C. Synchronization of the Time Signals

Now, the acoustic signal can be used to synchronize the devices. After a fixed time interval the acoustic signal is recorded, the devices are synchronized and thus the total error of time deviation can be reset to zero. We used smartphone SP1 and select different synchronization intervals.

TABLE I. STATISTIC VALUES ABOUT THE ERRORS OF THE SINGLE SMARTPHONES (ELAPSED TIME WINDOW 15 MIN)

device	Sp1	Sp2	Sp3
mean total error (mte) [ms]	2.0111	7.1256	9.2700
standard deviation (std) [ms]	(± 4.1750)	(± 4.705)	(± 4.3541)
maximum total error [ms]	13	18	20
mean single error [ms]	0.0100	0.0111	0.0189
standard deviation [ms]	(± 4.6101)	(± 4.5342)	(± 4.7198)
maximum single error (mse) [ms]	8	7	12

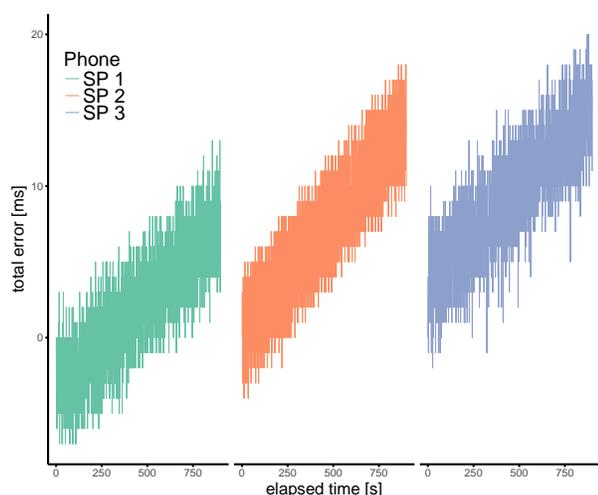


Fig. 9. Total error development of the three different smartphones

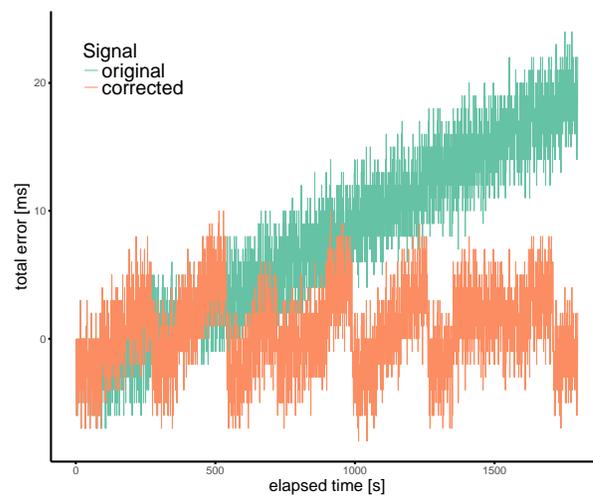


Fig. 10. Total error of the original signal and the corrected signal (adjustment interval: 90 ms)

In Table II the statistical values for the different adjustment intervals are given. It can be seen that the maximum total error, as well as the standard deviation, is nearly the same for all intervals. The mean total error increases with the longer interval lengths. Fig. 10 shows the two total errors for two different correction intervals, one from the original signal and one from the corrected signal. It can be seen that we can reduce the accumulation of the total error. To get a better understanding of the error distribution, in Fig. 11 the histograms of the total errors are visualized. For the original signal, the total error is widespread and has an overhang to larger positive values compared to the corrected signal. The corrected signal has smaller total errors and a smaller distribution around the zero point. The differences in the error distribution according to the size of the intervals are not decisive, as it can be seen in Table II.

D. Aligning Sensor Data based on a Time Synchronization

Assuming a situation in which a bridge has to be assessed. The assessment is based on the measured values of acceleration sensors built-in the smartphone. We use several smartphones in this situation since we can get a more accurate result. In order to align the measured values, these devices must be synchronized - the use case for our proposed approach of acoustic sound synchronization. The object of this experiment

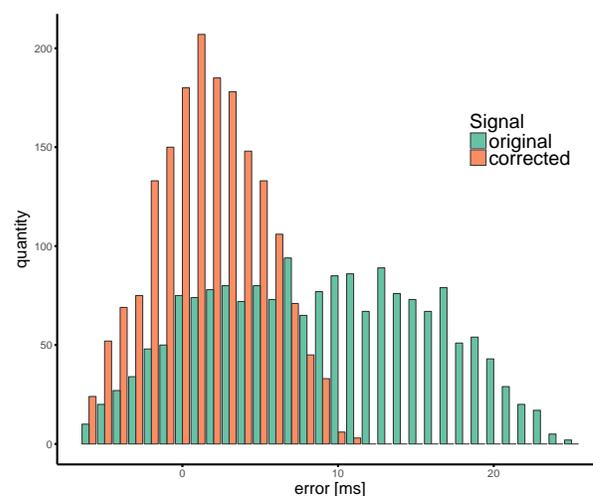


Fig. 11. Histogram of the total error distribution of the original signal and corrected signal (adjustment interval: 90 ms)

TABLE II. STATISTICAL VALUES ABOUT THE TOTAL ERROR FOR DIFFERENT ADJUSTMENT INTERVALS (MEAN TOTAL ERROR (MTE), STANDARD DEVIATION (STD), MAXIMUM OF ABSOLUTE ERROR (ME))

correction interval [s]	15	30	45	60	95	120	180	360
mte [ms]	0.095	-0.133	-0.3200	-0.113	0.620	-0.080	1.120	1.720
std [ms]	(±3.60)	(±3.60)	(±3.67)	(±3.37)	(±3.50)	(±3.18)	(±3.65)	(±3.52)
me [ms]	10	10	10	9	9	10	11	11

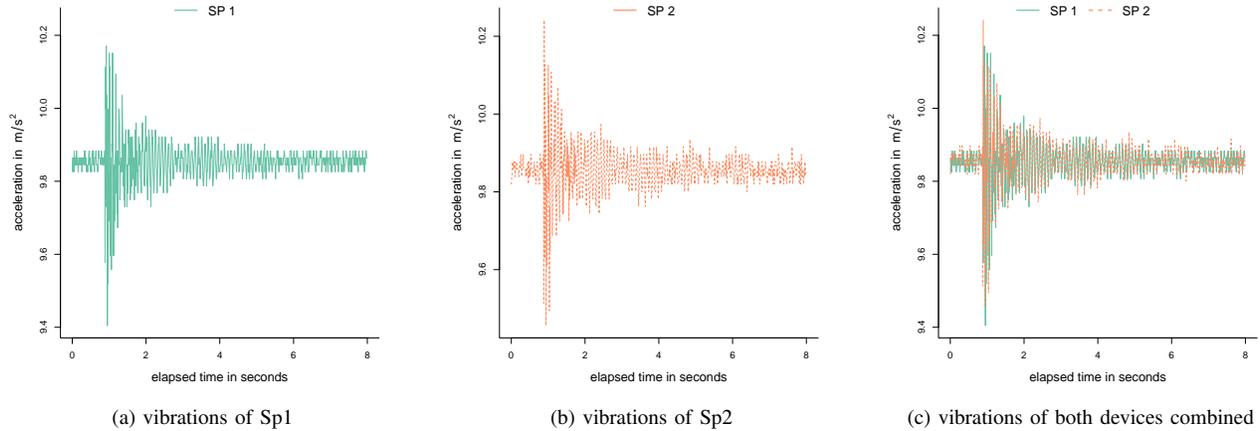


Fig. 12. Vibrations of a bridge triggered by a walking pedestrian and recorded by the smartphones. Fig. (a) shows the vibrations recorded by Sp1, (b) recorded by Sp2 and (c) the data of both devices aligned after time synchronization.

is a single span girder bridge (T-beam) having a structure of concrete-steel-composite (see Fig. 13). This bridge has a length of 30 meters, a width of 4 meters and weighs about 70 tons. During the experiment, the bridge was stimulated by a walking pedestrian. We placed the smartphones at about 1/6 of the bridge's length. Fig. 12a and Fig. 12b show the vibrations caused by the pedestrian and recorded by smartphone 1 and smartphone 2. As it can be seen, there are differences between the measured values of the two devices. Since the two devices were synchronized by our method, these measured values can be aligned. This leads to a more precise measurement result, as it can be seen in Fig. 12c and thus to a more reliable assessment of the accessibility of this bridge.

V. CONCLUSION

In this article, we introduced MobileTimeSync - an Android App for time synchronization on smartphones. This app



Fig. 13. The bridge for our experiment.

generates an audio signal that contains several peaks at the same distance from each other. Afterwards, this signal is sent out by the built-in loudspeaker of the device. This signal is received by all nearby devices and can be recognized by observing the amplitude of the built-in microphones and the corresponding timestamp. Finally, we achieved an accuracy of 10 ms. Thus, we confirm the results of Lazik et al. [16]. One reason for this value is the strong jittering on these devices. In the next steps, we have to investigate whether this accuracy is sufficient for the above-mentioned issue of damage localization. Furthermore, the jittering must be investigated. One option to avoid jittering could be the Native Development Kit (NDK), which allows native system calls.

Nevertheless, one advantage of this approach is that network, data or GPS connections are not necessary. The only limitation is the distance between the sound source and the devices. This means that this approach is highly suitable for areas where the above-mentioned synchronization services are not available. Rolling ahead, we want to bring this App to the Android Play store.

ACKNOWLEDGEMENT

This research has received funding from the German Federal Ministry of Education and Research within the project "Vulnerability of Transportation Structures – Warning and Evacuation in Case of Major Inland Flooding" "FloodEvac" (No. 13N13194/13N13195). Special thanks to Bharat Yadav for proofreading.

REFERENCES

- [1] M. Benndorf and T. Haenselman, "Time synchronization on android devices for mobile construction assessment," in *Tenth International Conference on Sensor Technologies and Applications SENSORCOMM 2016*. IARIA, 2016, pp. 83–87.
- [2] M. Benndorf, M. Garsch, T. Haenselmann, N. Gebekken, and I. Videkhina, "Mobile bridge integrity assessment," in *Proceedings: IEEE Sensors 2016*, 2016.
- [3] M. Benndorf *et al.*, "Robotic bridge statics assessment within strategic flood evacuation planning using low-cost sensors," in *Safety, Security and Rescue Robotics (SSRR), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 13–18.
- [4] K. Sun, P. Ning, and C. Wang, "Tinysersync: secure and resilient time synchronization in wireless sensor networks," in *Proceedings of the 13th ACM conference on Computer and communications security*. ACM, 2006, pp. 264–277.
- [5] B. Kusy *et al.*, "Elapsed time on arrival: a simple and versatile primitive for canonical time synchronisation services," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 1, no. 4, pp. 239–251, 2006.
- [6] J. Postel, "Rfc 768 - user datagram protocol," *Isi*, 1980.
- [7] D. L. Mills, "Network time protocol (ntp)," *Network*, 1985.
- [8] —, "Internet time synchronization: the network time protocol," *Communications, IEEE Transactions on*, vol. 39, no. 10, pp. 1482–1493, 1991.
- [9] J. Levine, "Timing in telecommunications networks," *Metrologia*, vol. 48, no. 4, p. S203, 2011.
- [10] A. N. Novick and M. A. Lombardi, "Practical limitations of ntp time transfer," in *Frequency Control Symposium & the European Frequency and Time Forum (FCS), 2015 Joint Conference of the IEEE International*. IEEE, 2015, pp. 570–574.
- [11] K. Zhao, A. Zhang, and K. Liang, "Research on the uncertainty evaluation of network time service system," in *European Frequency and Time Forum (EFTF), 2012*. IEEE, 2012, pp. 122–125.
- [12] S. Ping, "Delay measurement time synchronization for wireless sensor networks," *Intel Research Berkeley Lab*, 2003.
- [13] W. Commons, "Ntp— file: -jpg," 2016, [Online; accessed 01-March-2018]. [Online]. Available: <http://bit.ly/2F04ILs>
- [14] D. C. Mazur, R. A. Entzminger, J. A. Kay, and P. A. Morell, "Time synchronization mechanisms for the industrial marketplace," in *Industrial & Commercial Power Systems Technical Conference (I&CPS), 2015 IEEE/IAS 51st*. IEEE, 2015, pp. 1–7.
- [15] S. Sridhar, P. Misra, G. S. Gill, and J. Warrior, "Cheepsync: a time synchronization service for resource constrained bluetooth le advertisers," *IEEE Communications Magazine*, vol. 54, no. 1, pp. 136–143, 2016.
- [16] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Ultrasonic time synchronization and ranging on smartphones," in *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2015 IEEE*. IEEE, 2015, pp. 108–118.
- [17] S. Giucastro, "Getting high precision timing on android," 2012, (retrieved: 13-Mar-2017). [Online]. Available: <http://ubm.io/2HWdiz>
- [18] Google, "Android developers systemclock," (retrieved: 15-Mar-2017). [Online]. Available: <https://developer.android.com/reference/android/os/SystemClock.html>
- [19] K. Gopal, "Truetime for android," 2016, (retrieved: 29-Mar-2017). [Online]. Available: <https://github.com/instacart/truetime-android>
- [20] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: a survey," *Network, IEEE*, vol. 18, no. 4, pp. 45–50, 2004.
- [21] T. Haenselmann, *Wireless Sensor Networks: Design Principles for Scattered Systems*. Oldenbourg Verlag, 2011.
- [22] J. Aron, "Flickering lights could help smartphones keep time," *New Scientist*, vol. 215, no. 2874, p. 22, 2012.
- [23] Z. Li *et al.*, "Flight: Clock calibration using fluorescent lighting," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 329–340.
- [24] E. Peguero, M. Labrador, and B. Cook, "Assessing jitter in sensor time series from android mobile devices," in *Smart Computing (SMART-COMP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–8.

Evaluating Key Factors Influencing ERTMS Risk Assessment: a Reference Model

Katja Schuitemaker
Design, Production and
Management Department,
University of Twente
Enschede, the Netherlands
e-mail:
k.schuitemaker@utwente.nl

Heidi van Spaandonk
Safety Department,
ProRail
Utrecht, the Netherlands
e-mail:
heidi.vanspaandonk@prorail.
nl

Marco Kuijsten
Safety Department,
NS,
Utrecht, the Netherlands
e-mail: marco.kuijsten@ns.nl

Mohammad Rajabalinejad
Design, Production and
Management Department,
University of Twente
Enschede, the Netherlands
e-mail:
m.rajabalinejad@utwente.nl

Abstract— The European Railway Traffic Management system (ERTMS) aims to replace the various national train command and control systems in Europe, and will serve to improve cross-border interoperability, with the final aim of improving the competitiveness of the rail sector. As an additional effect, it is argued that implementation of ERTMS will improve safety. To provide insight into safety developments within the European railway system, this study evaluates ERTMS at both the national and international levels. For this purpose, international data from European ERTMS implementations is combined using data obtained from interviews with ERTMS stakeholders and safety experts from the Netherlands. Effects of the safety case regime, interoperability, deregulation and dynamic specifications on the European railway system have been researched. We present our findings into a reference model that describes the existing situation and shows what key factors are most suitable to improve the situation. The challenges are to improve resilience, to generate more awareness of interrelationships between hazards and risks, but even more: comprehending the safety architecture and creating cross-discipline understanding.

Keywords - ERTMS; railway safety; interoperability; risk management.

I. INTRODUCTION

This paper is an extension to the work presented at the Seventh International Conference on Performance, Safety and Robustness in Complex Systems and Applications (PESARO) 2017 conference [1]. Safety has always been one of the key priorities in the railway industry. There have been many initiatives to improve railway safety in the Netherlands, some of them which are listed below:

- Reducing the number of train-on-train collisions by, among other actions:
 - Implementation of the Routelint system: providing train drivers with real-time information about nearby trains.
 - Implementation of the improved version of the Dutch Automatic Train Control System (ATB Vv).
- It has been argued that the implementation of the European Railway Traffic Management System (ERTMS) has beneficial effects on the overall safety level of the railway system [2]. As ERTMS seems to hold great promise at both the Dutch national and international levels, we will discuss this system in more detail.

The European Union (EU) has adopted Directives concerning the interoperability of the European railway system and railway safety. The implementation of these Directives is aided by committees such as the European Committee for Electrotechnical Standardization (CENELEC). The CENELEC European Norm (EN) 5012x is a family of standards that contain requirements and recommendations concerning processes to be followed for the development and assurance of safety-critical systems. As part of CENELEC, the EN50126 (The Specification and Demonstration of Reliability, Availability, Maintainability and Safety) describes a performance-based approach that includes proactive argumentation on why a system is acceptably safe. The railway industry in the Netherlands tends to follow the safety case regime recommended in EN50126. As early as the 1990s, the European Commission (EC) decided that passenger trains should be able to travel seamlessly across international borders in Europe. In 1998, the EC requested the foundation of the Union Industry of Signalling (UNISIG) and assigned this with the task of drafting the technical specifications for ERTMS. The International Union of Railways states that the goal of ERTMS is “to enhance cross-border interoperability and signalling procurement by creating a single Europe-wide standard for railways with the final aim of improving competitiveness of the rail sector” [3]. ERTMS is a command, control, signalling, and communication system for railway management and safe regulation. It is composed of two technical layers:

- European Train Control System (ETCS): the Automatic Train Protection (ATP) system that makes sure trains do not exceed safe speeds or run too close together.
- Global System for Mobile Communications – Railways (GSM-R): helps provide communication for voice and data services.

Beyond these two technical layers are the European Operating Rules (EOR) and European Traffic Management Layer (ETML).

- ERTMS implementation can vary in:
- Level: basic concepts of the ERTMS.
 - Baseline: corresponds to the version of the technical specifications.
 - Version: modification of properties.
 - Operation mode: various conditions required for managing various situations.

ERTMS has become the European standard for the Automatic Train Control (ATC) that facilitates an interoperable railway system in Europe. This means that ERTMS allows trains to run across borders without changing their control systems. Recently, the ERTMS European Deployment Plan (EDP) has set targets that by 2023 50% of the core network corridors will be equipped with ERTMS. The aim is to implement ERTMS Level 2 on key routes in the Netherlands by 2028. With these aims in mind, there is the promise of an increase of railway safety by implementation of ERTMS, which self-evidently would be very beneficial for railway industry and the general public. Some reasons why ERTMS is considered to increase railway safety include:

- ETCS supervises both the position and speed of trains to make sure they continuously remain within the allowed speed and distance limits, and, if necessary, it will command the intervention of the braking system to avoid any collision [2]. The train can continually receive authorisation to continue running at maximum allowed speeds through the GSM-R system (only available at ERTMS level 2 and when driving in the correct operation mode).
- Reduce the risk of human error [4]; for example, work-related errors caused by stress, tiredness, fatigue, and sleep disturbance.
- Decreasing the number of Signals Passed At Danger (SPADs) [5][6]. This can be explained by the fact that ETCS is able to intervene in the braking curve for a train driving at any speed.

However, there is evidence that implementation of ERTMS does not automatically mean a safer railway system. For example, in practice, in the Netherlands, SPADs still occur when driving with ERTMS [7], see Table I.

TABLE I. SPADs UNDER ERTMS

	2013	2014	2015	2016
SPADs under ERTMS Level 1	1	0	1	4
SPADs under ERTMS Level 2	9	10	16	15

However, these numbers are low, so they can only be considered as an indication, and not necessarily representatives of a trend. More studies, both scientific and industrial, question the safety benefits arising from the implementation of ERTMS.

At the international level:

- Smith et al. addresses issues relevant to the safe introduction of ERTMS into European railway systems [8]. These issues include technical system integration, technical system failures and human factor considerations.
- Laroche and Guih ery study the European Transport Policy, the role played by the EC, and the ERTMS innovation process in relation to innovation processes in surface transport, and the difficulties inherent in the

implementation of an intelligent transportation system innovation [9].

- Ghazel addresses the regular evolving documents that give rise to successive ERTMS versions [10].
- The EC itself has studied past and current problems resulting from ERTMS implementation [11].

At the Dutch national level:

- The Ministry of Infrastructure and the Environment, ProRail and NS have collected information on ERTMS standard usage for safety systems, and the effects on various goals of the Railmap 1.0 [12].
- ProRail and NS executed a pilot to gain experience with driving under ERTMS [13].
- A specialised team investigated the sequence of events and decision-making processes in the Netherlands that have led to delays in deployment of the ERTMS train signalling systems in the High Speed Line (HSL) railway project [14].

This study evaluates the effects of the safety case regime, the inclusion of various ERTMS specifications, the exclusion of a responsible integrator, deregulation, and the final effects for risk assessment and safety. The focus is on risk management and safety of ERTMS as a System of Systems (SoS). We bring all our findings together and represent them into a reference model that illustrates the main line of argumentation. This reference model points out the key factors that are most suitable to address in order to improve the situation.

Section II provides an overview of the background of railway deregulation, ERTMS specifications, European interoperability, and cost reduction that are the results of changes in organisational behaviour. The methodology is discussed in Section III. Section IV explains findings with regard to the number of stakeholders and accompanying views, interests and interactions, possible local goal trade-offs, various interpretations, the decision-making processes, borders between disciplines leading to unique design realisations, and the relationship between overview and safety architecture understanding. Findings are discussed in Section V. Section VI summarises the findings, draws conclusions and highlights challenges.

II. BACKGROUND

Every railway system faces technical, managerial, organisational, and regulatory challenges. The subsystems can work perfectly individually, but together they can create a hazardous state. Many factors, both technical and socio-institutional, need to be combined to turn the challenge of one European train system into a great success, satisfying social needs for lower costs, better utilisation of infrastructure, and less complex logistics [15]. During this study, several developments in the rail industry appeared to have a great effect on the safety level of a railway system including ERTMS.

A. Deregulation

Regulation is required to prevent monopoly exploitation, to reduce asymmetry of information, to guarantee non-discriminatory access to any essential facilities and to monitor the performance of a service provider [16].

Deregulation is the reduction or elimination of government power in an industry, usually promoted to encourage more competition within a specific market. Starting in the 1990s, in order to promote greater competition, the rail industry in Europe has been gradually restructured. On the one hand, vertical separation means that the management and ownership of infrastructure are totally separated from the operation of passenger and freight rail services. On the other hand, multiple operators are using the infrastructure. In some countries, infrastructure has been separated from train operations, whereas in others, this has not been the case. Privatisation and deregulation have led to an increased involvement of private actors, both nationally and internationally [17].

B. ERTMS specifications

The Union Industry of Signalling (UNISIG) was founded in 1998/99 at the specific instigation of the European Commission (EC) [18]. It was created to develop ERTMS specifications. The final version of ERTMS specifications was published by the EC following the approval of the Member States. In November 2012, the EC intentionally deleted ERTMS Functional Requirement Specifications, making these specifications no longer mandatory. The remaining System Requirements Specifications are written in a natural language. These specifications allow multiple interpretations [19].

C. Interoperability

The meaning of interoperability is two-fold. On the one hand, interoperability refers to a geographical interoperability among countries and among projects. On the other hand, it also refers to interoperability among suppliers. This opens the supply market and increases competition within the industry [20]. The result of this is the absence of a single entity that is responsible for the railway system as a whole.

D. Safety case regime

The safety case approach is goal-oriented, meaning that organisations should always seek for improvements in safety. It requires a detailed hazard analysis comprising causal analysis, a dedicated hazard identification focusing on the system under consideration, and a common cause analysis. A hazard is a potential source of harm in a system. Hazard identification is performed during risk assessment and within hazard control. The results of risk assessment are a set of safety requirements that define the required efficiency of safety functions. These can be assessed both quantitatively and qualitatively. Taken into consideration the identified hazards, risks should be reduced to an acceptable level. It has to be demonstrated that the risk is reduced to ‘As Low As

Reasonably Practicable’ (ALARP). In the end, one is able to argue whether the system is acceptably safe.

E. Cutting cost and time

According to Rasmussen, systems and organisations continually experience change, as adaptations are made in response to local pressures and short-term productivity and cost goals. Several accidents such as Bhopal, Flixborough, Zeebrugge, and Chernobyl demonstrate that they have not been caused by a coincidence of independent failures and human errors, but by a systematic deterioration of organisational behaviour towards an accident under the influence of pressure towards cost-effectiveness in an aggressive, competitive environment [21].

In order to reduce the risk, Dutch national safety goals are approached through use of the ALARP-principle and standstill-principle. This means that all risks must be reduced such that they are below a threshold of practicability. For risks in the “ALARP area”, all potential risk reducing measures must be evaluated in terms of cost efficiency, cost-benefit balance or some similar economic measure. Selected risk-reducing measures may be introduced based on experience or best practice in combination with cost-efficiency considerations [22].

III. METHOD

The objective of this empirical research is to identify key factors and interrelationships of the safety of ERTMS. Emphasis was placed on the ERTMS safety architecture, and the relationships between social and technical safety entities of ERTMS at both the Dutch national and international levels. The findings in this paper are based on international data from European ERTMS implementations, linked with national data obtained from semi-structured interviews based on questionnaires.

To investigate the nature of phenomena, we adopt qualitative and quantitative analysis methods in the form of standardised interviews. We used an interview guide with a list of questions generated in advance, allowing the same topics to be covered during interviews, and at the same time, if necessary, leaving room for more exploration of certain issues. This interview guide consisted of a brief description on one’s background and relationship with ERTMS, open questions with regard to pros and cons of ERTMS, and more detailed question to cover specific topics. For this research, the data are transcribed and analysed using Grounded Theory data analysis [23]. By constant comparison, every new piece of data is compared with earlier data to find similarities and differences. The data was used for formulation of hypotheses, and hypotheses are verified logically on internal validity and external validity.

The topics discussed include the effects observed from inclusion of various ERTMS specifications, the exclusion of a responsible integrator, deregulation, and the final effects for risk assessment and safety. Systems under consideration are ERTMS Level 1 and ERTMS Level 2. Projects discussed during interviews are the five ERTMS-projects in the

Netherlands (Betuweroute, Port of Rotterdam, HSL South, Amsterdam-Utrecht and Lelystad-Zwolle).

Participants were Dutch ERTMS key stakeholders and safety experts from the Ministry of Infrastructure and the Environment, train operating companies, infrastructure managers, and independent consultants involved with the ERTMS national program, each representing their own viewpoint. For increasing validity and minimising subjectivity, we ensured each topic was discussed from various points of view (political, company, management, operations). The average number of years of experience of the participants varies from 2 to 14, with an average of 8. The educational background was mostly technical or safety related, with a few exceptions. Participants are informed about the aim of this study beforehand the interview. Also, it was explained that they should have in mind their own expertise when answering the questions, meaning, they should respond from their own viewpoint, not from someone else's viewpoint. Contact between researcher and participants has been direct.

Interviews lasted between 30 and 90 min. Data was collected between February 2016 and August 2016. In total, 15 semi-structured interviews were held, performed face to face. All interviews were audio recorded, transcribed verbatim, and summarised. Transcriptions were processed through qualitative inductive content analysis in order to develop a theory, and identify themes through repeated examination, comparison, abstraction, and data reduction. The material was abstracted and reduced to a set of themes. Resulting themes were quantified and integrated with the responses. The procedure was repeated to refine chosen themes. Two main categories were identified as a thread through transcriptions: (1) implications with regard to socio-technical safety; and (2) implications with regard to the safety architecture. Using the Design Research Methodology [24] as a supporting framework, key factors found were translated into a reference model, which graphically shows the current understanding of the safety challenges of ERTMS.

As for verification, summaries were sent to the interviewee. Most interviewees made small corrections in the summaries. The reference model is logically verified by consistency, meaning there are no internal conflicts between interview answers, key factors and well-established literature. For this literature review, the high level goal is to identify supporting evidence and contradictions. For the search string:

Part 1

The first part captures keywords related to the system under consideration such as “*European Railway Traffic Management System*”, “*ERTMS*” or “*ETCS*”.

AND

Part 2

The second part captures keywords related to safety such as “*railway safety*”, “*safety case*”, “*safety analysis*”, “*risk assessment*”, “*risk analysis*” or “*risk evaluation*”.

OR

Part 3

The third part captures keywords related to complexity and interdisciplinary such as “*architecture*”, “*socio-*

technical safety”, “*decision-making process*”, “*integral assessment*”.

Next to this, we also verify logically on internal validity with the meaning of causes and effects of key factors that could be interchangeable, and external validity with the meaning of participant- and time dependency.

IV. FINDINGS

To obtain an understanding of the existing situation, we represent our findings by the creation of a reference model. This reference model consists of key factors and links among key factors that can come from resources, assumptions, or the experiences of stakeholders. In the end, many key factors influence safety of ERTMS, and result in a high complexity. A graphical representation is created to provide an overview. In order to limit the amount of information, this representation is divided into two parts: a lower part, and an upper part.

A. Reference model lower part

Next, we discuss the lower part of the reference model, representing implications with regard to socio-technical safety. Key factors identified include effects of the safety case regime and ERTMS specifications, and how the decision-making processes and missing integrator influence risk assessment.

1) Safety case regime

The safety case is the documented demonstration that the product, system or process complies with the relevant safety requirements. It can be seen as a risk- or hazard management framework, where the organisation identifies controls to deal with identified hazards and measures. Such controls must ensure the continued working of safety-related functions. The Netherlands evaluates risks using the ALARP-principle, described by EN50126, allowing cost-effectiveness of safety measures to be explicitly considered. In order to classify risks, hazards are categorised on frequency and severity, resulting in a risk matrix. High risks require mitigation or risk acceptance. The European Union Agency for Railways explains a shift from quantitative data to qualitative data [25]. As a result, organisations must come up with descriptions instead of numbers and observe data rather than just collecting data. This qualitative data is based on the logical reasoning of many experts. Stakeholders experience this regime as a challenge, as logical reasoning can cause multiple interpretations. Also, they experience the safety case approach as complex and time-consuming. For example, in practice, the safety case for the High Speed Line South resulted in only addressing major hazards due to time pressure [26].

Barua explains that one disadvantage of the safety case approach is that the explanation and interpretation of the desired performance levels expressed in the regulation can be both complex and challenging [27]. Nair identifies 25 studies citing ambiguities in the application of standards, such as the existence of multiple interpretations of the evidence requirements in the standards [28].

According to the ERTMS strategy group in Great Britain, initially the principal motivation for ERTMS was to further improve safety. “Over approximately the last ten years, capacity became a more significant influence and then, more recently, cost reduction [29].” Demonstration of compliance by reference to safety standards is usually costly and time-consuming [30]. Pressure towards cost-effectiveness can inadvertently lead to generating adaptive responses. According to Leveson [31], pressure towards cost-effectiveness and increased productivity is the dominant element in decision making.

2) ERTMS specifications

ERTMS is not fully specified. Rather it is a system architecture, which describes how a range of elements should function. Earlier studies explained that ERTMS specifications are unstable [9][32], written in informal language [10], non-consolidated [8] and incomplete [15][33]. To be more specific, missing parts concern management, integral system integration, and physical design. Therefore, stakeholders still claim that the specifications are not sufficient. These specification deficiencies have effects on system safety in the three ways described below.

- The management of railway signalling in ERTMS is not based on global rules, meaning they are customised to each country. Every country uses its own interpretation of the European Norm (EN) 50126. It is therefore difficult to compare the systems in terms of safety.
- As for system integration, the lack of harmonised specifications requires the development of various solutions for each project. As for safety, this implies a need for additional safety analyses.
- Open specifications that exclude physical design, result in unique ERTMS design realisations, and the occurrence of not fully compatible solutions for each system to be developed. Again, this implies a need for additional safety analyses.

In practice, updates to new systems are postponed in anticipation of new specifications, covering multiple requirements through one update. Experts explain that unstable specifications make it difficult to adopt innovations and that problems occur with adapting the new system to the old one. There is consensus among experts about the effects of the inclusion of various ERTMS specifications on the safety of the System of Systems (SoS). 93% of the participants explained the relationships between a high variety of ERTMS specifications and the decision-making process.

3) Decision-making

Today’s ERTMS requires strategic safety decisions concerning functionality and policies. This process is based on criteria resulting from risk analyses, and organised through a specified organisation. Experts explain the open specifications complicate the architecture, but also safety decision-making processes. As a consequence of incompleteness, it could be possible to make a decision based on implicit factors. For example, as safety is sometimes seen as hindrance to effective marketing solutions, the focus can be on finishing on time and

approving a design. In the case of a complex architecture, it can be difficult to identify hazards and what control measures must be taken by whom. Care should be taken that risk is evaluated using complete information.

According to the parliamentary commission that looked at the failed Fyra project, it seemed that safety had become a subject for negotiation [34]. As also described by Nusser [35], “Black box” approaches are regarded with suspicion – even if they show a very high accuracy of available data – because it is not feasible to prove that they will show a good performance under all possible input combinations.

Late safety inclusion is also questioned by Enserink in the context of major projects in the Netherlands: “It is strange to see how in many large projects, such as the Westerschelde tunnel, the Betuwelijn, and the ‘Groene Hart’ bored tunnel for the High Speed Line South, the discussions of safety issues and safety management took place at a very late stage in the project cycle” [36]. He also explains: “In all the examples in the planning phase, the analysts neglected the safety issues or these issues were temporarily stalled because of their complexity.” According to Høj et al.: “Biased results may result from excluding certain events, different analytical methods and data, different modelling assumptions, different methods for including uncertainties, best estimates vs. estimates “on the safe side”, etc.” [37].

4) Integrator

EN50126 describes a simplified approach that helps organisations to conduct risk assessment and hazard control. This approach, the so-called hourglass model, provides an overview of the major safety-related activities that are required to reach an acceptable safety level for a technical system, including defining the corresponding responsibility areas. This hourglass model is shown in Fig. 1. This model provides an overview of major safety-related activities that are needed to ensure an acceptable safety level for a technical system, including the corresponding responsibility areas. As shown, responsibilities overlap. According to EN50126, risk assessment should be done at the railway system level, where it relies on system definition, and includes risk analysis and risk evaluation. This risk analysis includes hazard identification, consequence analysis, and the selection of risk acceptance criteria. Hazard analysis at the level of the system under consideration includes a causal analysis, hazard identification focusing on the system under consideration, and a common cause analysis. CENELEC requires that, during each project, responsibilities have to be clarified unambiguously in order to avoid gaps or overlaps.

The Dutch ERTMS program is a collaboration of stakeholders, not as if it were a single person that acts for certain purposes. Experts explain that without the ERTMS program being a legal entity, it is difficult to allocate hazards and risks, especially the ones on interfaces, to a responsible stakeholder. As for the Netherlands, what makes this even more difficult is a missing central designer, or indeed any party, that knows the entire complex system. Although the Common Safety Method (CSM) aims at an integral safety approach, the final report on the ERTMS pilot between Amsterdam and Utrecht explains that “overarching processes

between railway and train transportation are missing and that these are necessary for optimum implementation of ERTMS” [13]. One of the safety case principles is that those who create risks are responsible for controlling those risks. Experts explain that as a result, organisations feel responsible for their own processes, not for the integral railway system.

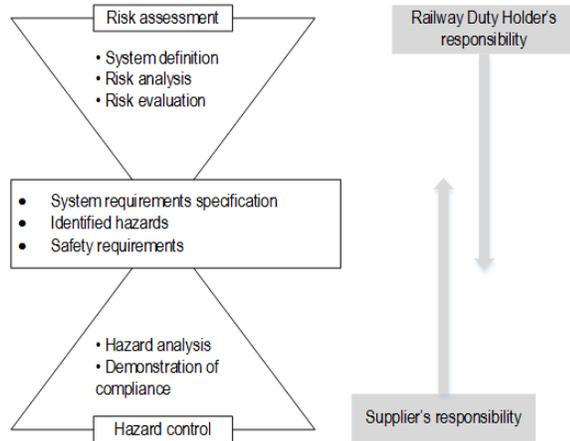


Figure 1. Hourglass model described by EN50126.

Lack of responsibility is also recognised by the Dutch Ministry of Infrastructure and the Environment [13]. This is a known challenge explained by Harvey: “Complexities at the organisational level can breed a culture in which no-one is willing to accept responsibility for risk and blame is always shifted to a higher level in the network [38].”

Next to the hourglass model, safety activities tend to be performed in parallel with systems engineering activities [39] [40]; see Fig. 2. Shown is the left side of the V-model [41]. Squares on the right represent well-known stages in system development. Squares on the left represent safety activities executed during each of the system development stages. For example, safety engineers analyse requirements for hazards and communicate these to system designers. According to Mauborgne et al., who defined a missing link between safety and systems engineering, the two activities are not always well integrated [42]. Experts explain communication of safety integration in the development phase as a challenge, especially with regard to defining interfaces.

Communication both within development teams and between individual developers is considered to be a source of safety-related faults in critical systems [43] [44]. According to Leveson, the defence community tried using the standard safety engineering techniques on their complex new systems, but the limitations became clear when interface and component interaction problems went unnoticed until it was too late, resulting in many losses and near misses [31].

5) Risk assessment

The generation of adaptive responses, complexity in decision-making processes and a missing overarching view affect risk assessment in the three ways discussed below:

- First, adaptive responses can depend on a number of assumptions. Such assumptions can be explicitly

formulated, but the danger is in including these without being subject to uncertainty. Describing and making judgements about risks in advance can misguide risk assessment.

- Second, the decision-making process and risks assessment process are not always interrelated, resulting in a missing integral assessment in the Netherlands.
- Third, risk assessment must be performed at the railway system level. Experts indicate the challenge lies in the incorporation at the System of Systems (SoS) level, which requires an overview of all safety entities to be considered.

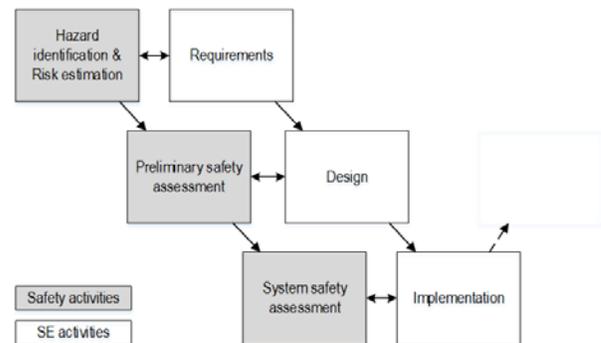


Figure 2. Relationships between safety and SE activities [38][39].

For example, misjudgements happened in relationship to the Piper Alpha accident in 1988: a series of events considered of low risk prior to the accident. Adaptive responses lead ultimately to related problems, having unforeseen ramifications in addition to system complexity, and therefore, to further unexpected behaviour, which reduces system safety [45]. At the technical level, much attention is given to safety issues, but hardly any attention is paid to safety at the level of decision-making [36]. According to Aven, the benefits related to the activity studied, as well as strategic and political aspects could be decisive for the decision making but may not be captured by risk assessment [46]. According to Harvey, the many and varied interactions among the individual components of a SoS produce emergent behaviour, which cannot be predicted on the basis of the performance of the individual subsystems in isolation [38].

The described relationships are shown in the lower part of the reference model of ERTMS in Fig. 3. This reference model represents the existing situation on implications with regard to the socio-technical safety of ERTMS. All ellipses represent key factors meaning those influencing factors that are considered as the core factors or root causes. The links describe the existing relationships. Not all relationships come from direct experiences. Therefore, these relationships are indicated as an assumption. The existing qualitative value of a key factor is represented by means of a ‘+’ or ‘-’. For example, the ‘+’ next to ‘number of different ERTMS inclusions’ indicates a high degree. Quantitative classification of whether or not participants experience a challenge in one of the key factors is shown in Fig. 4.

B. Reference model upper part

Next, we discuss the upper part of the reference model including implications with regard to the safety architecture.

Key factors identified include the effects of the deregulation and the increased number of stakeholders with their own specific languages and cultures. Also, how dynamic specifications have resulted in unique design realisations. We argue how the quality of overview influences risk assessment, and therefore, the overall safety level of ERTMS. Some key factors are safety related items that fundamentally contribute to a successful system operation [47].

1) Deregulation

Available historical data on fatal railway accidents has shown a solid gradual improvement in railway safety over the past three decades. However, this trend has slowed down since the late 1990s [48]. Increasing regulation, standardisation and systematisation have paid safety dividends, although an adverse effect is the increase in regulation. Safety regulation has increased a hundred-fold between 1947 and 2008 [49]. Experts indicate the deregulated organisation results in many stakeholders. As is also concluded by Iglesias [50], incomplete/unstable specifications of ERTMS are further hampered by company specific requirements.

2) Number of stakeholders

At the national level, the change from one national actor to multiple commercial actors shows an increase in operators (CFL, NS, SNCB, etc.) using ERTMS tracks. Also, with previous Dutch automatic train protection (ATB) tracks, only one manufacturer (Alstom) was involved. With ERTMS and the tendering of subsystems, various manufacturers (Alstom, Ansaldo, Bombardier, etc.) are involved. Stakeholders within the ERTMS program include train operating companies, infrastructure provider and independent consultants. At the international level, infrastructure managers (Deutsche Bahn, INFRADEL, ProRail, etc.) from various countries must collaborate in order to provide a seamless transition.

In the first place, the rising number of parties involved entails a considerable diversity of points of view, skills, responsibilities, and interests. Experts explain the challenge lies with the many interests, creating the risk of compromising too much on safety. For example, the train derailment in Hilversum shows that the commercial character of maintenance, specifically the introduction of market forces, has led to an unavoidable interplay of forces. According to the Dutch safety board: "In this context, the train derailment in Hilversum teaches us that the related interests can gradually and unnoticed apply pressure on the management of safety risks [51]".

3) Language and culture

Laurino explains that the historical world-wide railway framework is modified to country specific approaches to

public policy, geographical context, transport system, economic situation, business and regulatory environment [52]. On top of this, countries use their own language, making intersectional challenges even more complex. Somerville describes that much of the work of professionals is knowledge-based and reflects their professional discipline, training and culture [53]. As an example, experts explain differences in both language and culture can lead to different people doing the same job, but working in different ways, leading to differences in understanding the overall risk assessment and evaluation. For this reason, risk perception can be different per organisation, even different per stakeholder.

For these reasons, at both the international level and national levels, stakeholders experience difficulties with understanding their respective systems. This is also described by Forsberg [4], who states that the new societal organisation, where rail transport is controlled by an increasing number of mainly private actors, intersectional issues and decisions have increased among the various actors. This happens particularly since mishaps or accidents are often caused by circumstances or weak links between them.

4) Boundaries

Experts explain the increased number of stakeholder viewpoints in ERTMS result in a complex architecture. They explain a system becomes complex when it is composed of many components that interact with each other. Complex intersections affect clarity of boundaries among subsystems within the system and between the system and its interaction with the environment.

Dekker explains that with increasing complexity, boundaries of what constitutes the system become fuzzy; interdependencies and interactions multiply and mushroom [54]. According to Rasmussen [21], people under pressure tend to explore and sometimes cross boundaries of safe operations.

5) ERTMS specifications

With the gradual implementation of a single signalling system through the EU, the EC has opted for radical innovation for all Member States. In the same vein, the Netherlands has opted for innovation in the form of a systems leap from traditional ATB to ERTMS. This can be contrasted with, for example, Belgium, which has opted for a more incremental development. Preferences vary at both the international and national levels. The signalling system for the Netherlands – Germany trajectories (remote monitoring) differs significantly from the signalling system for the Netherlands – Belgium – France trajectories (more autonomy for the train driver), which is more in line with ERTMS Level 2. Therefore, to migrate to ERTMS Level 2, France does not have to change much. To migrate to ERTMS Level 2, Germany and the Netherlands face a break with the past. Both the signalling systems and the automatic train protection systems are still markedly different from one EU country to another [55]. In addition, the various ERTMS levels include varying technical requirements and applications.

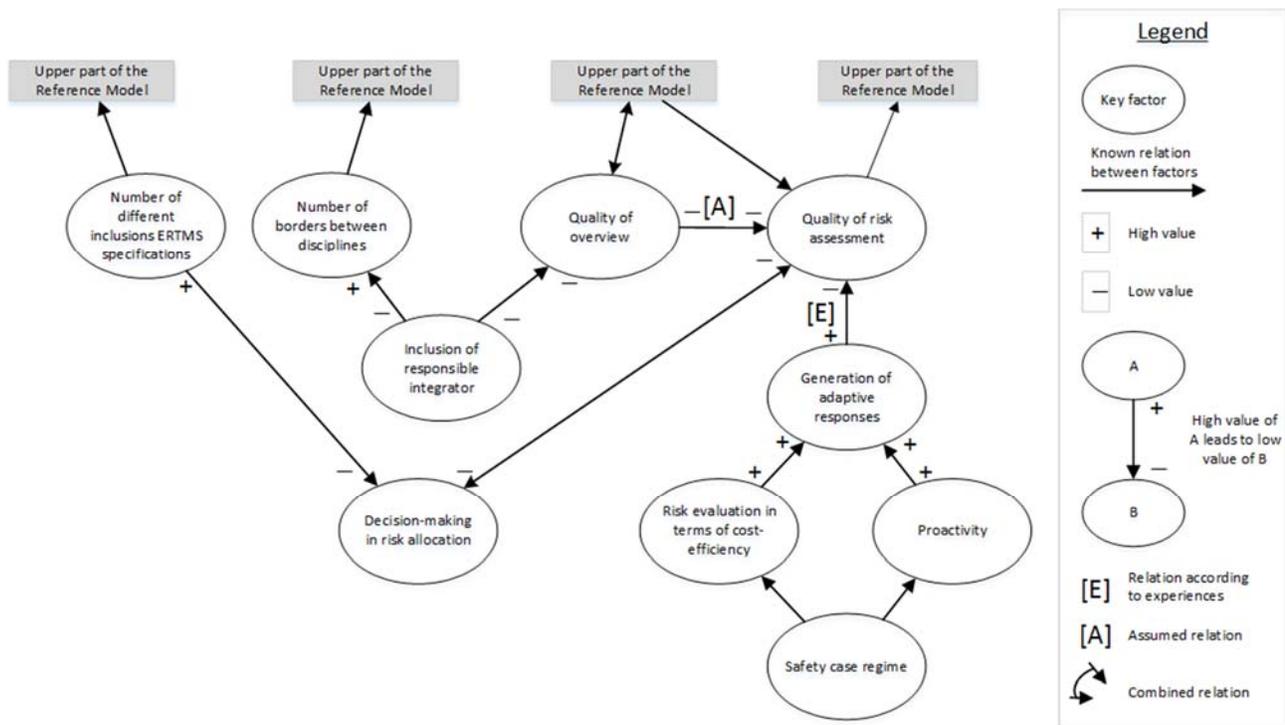


Figure 3. Lower part of the reference model on implications with regard to socio-technical safety of ERTMS.

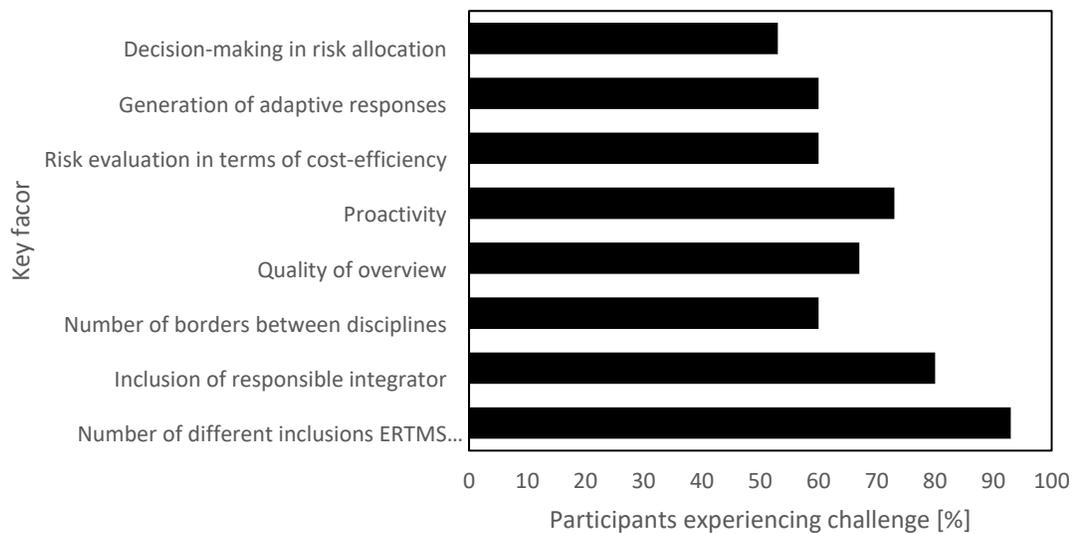


Figure 4. Participants experiencing challenges in key factors.

Various ERTMS subsystems have been tendered. In practice, the various stakeholders, even at the Netherlands national level where there have been 5 ERTMS-projects so far (see Table II), every subsystem is considered unique, and requires customised procedures and processes. Next to the 5 ERTMS-projects, Table II also shows which manufacturer was contracted, what ERTMS level was implemented, and when the track became operable.

TABLE II. ERTMS-PROJECTS IN THE NETHERLANDS

Project	Supplier	ERTMS level	In service date
Betuweroute	Alstom	2	2007
Port Rotterdam	Alstom	1	2009
High-Speed Line South	Thales/Siemens	1 and 2	2009
Lelystad-Zwolle railway	Alstom	2	2012
Amsterdam-Utrecht railway	Bombardier	1 and 2	2013

The various levels have been defined to allow each individual railway administration to select the appropriate ERTMS/ETCS application trackside, according to their strategies, to complement their trackside infrastructure and to achieve the required performance [57]. Considering only permitted disparities, such as varying ERTMS levels (0, STM, 1, 1 infill, 2, 3), already result in 31 possible transitions; see Table III.

TABLE III. ERTMS POSSIBLE TRANSITIONS

To From	0	STM	1	1 infill	2	3
0						
STM						
1						
1 infill						
2						
3						

Once a hazard scenario is identified, it is not trivial to identify all the possible causes in the system [33]. In other words, a system that is new, or particularly complex, can generate scenarios that are not generated during hazard analysis of previous comparable systems. Stoop also comments on the underestimation of ERTMS development: "There has been tension between incremental progress on the one hand and implementation in an existing railway network on the other hand with the ambitions on innovative ERTMS and public-private partnership" [14].

A higher level involves less side track equipment, but more on-board equipment. This change also implies that many of the costs of the signalling system will migrate from infrastructure managers to train operators. Infrastructure managers anticipating developments, whereas operators are reluctant to upgrade existing rolling stock [56], could be an explanation for the varying preferences.

6) Borders between disciplines

Experts explain employees are often focused on their own job. Ascribing meaning to data, so that it could be more readily used by others, is of less importance. Experts know a lot about their own subject, though knowledge of each component is limited. As a result, data is set and sent to the next disciplines.

This is also explained by literature. As described by Baxter [58], borders between disciplines have been largely maintained despite efforts at creating interdisciplinary teams by involving domain specialists in the design process. The success of system implementation is dependent on effective cross-discipline communication. The borders between disciplines are a known challenge for the safety of Systems of Systems [41].

7) Comprehending the architecture

Problems often arise at these borders due to a lack of shared understanding among the developers of subsystems. Gaps in assumed knowledge can influence both understanding and integration. The majority of the experts (87%) indicated difficulties with comprehending the architecture, and how this low comprehension affected the accomplishment of their tasks. One of the safety architects tried to obtain insight, by creating physical overviews of the SoS.

Leveson [59] states that lack of shared understanding among the developers of subsystems create coordination problems, ambiguity, and conflicts among independent decisions.

8) Number of unique design realisations

In practice, various preferences and implementations of ERTMS subsystems result in many transitions among various subsystems. In other words, implementation is unique for every project, and dependent on stakeholders, the environment, and activities. Table III shows just a fraction of the number of possible transitions. As a result, systems can be incompatible. For example, the two implementations made by Alcatel (Dutch part of the railway) and Alstom (Belgium part of the railway) that differed too much, so that a so-called gateway (network node) was necessary to transition from one system to another.

As is also explained by Leveson [57], the interconnectivity and interactivity among system components implies that greater complexity leads to vastly more possible interactions than could be planned, understood, anticipated or guarded against. As also concluded by Smith [8], the existence of many versions of ETCS with technical problems require a backup system.

9) Number of procedures and processes

In the end, when using ERTMS, the complexity of technology, use, and processes of the railway system increases. Experts indicated that the technological developments in ERTMS are underestimated. A failure when using ETCS can have up to 100 causes. Train drivers and signallers must find a solution through applying difficult procedures and processes, while using limited technical system knowledge. A large number of human resources

executing unique safety-critical procedures, increase the risk for human errors and therefore, influence the overall safety level of the system.

10) Local goal trade-offs

Stakeholders that are involved in multiple projects can have, depending on the meeting, various goals. Experts explain the creation of fuzzy boundaries and the lack of understanding of the safety architecture, allow local actors to change their conditions in one of its corners for good reasons, and without apparent consequences. This can bring immediate gains on some local goal trade-off. In other words, the decision-making process can be person-focussed, instead of organisation-focussed.

Both Leveson [59] and Dekker [54] explain that with a high number of widely distributed interacting components in an organisation, small 'drifts' in procedure or policy will not necessarily be identified as risks to the safety of the SoS.

11) Overview

As stated before, according to EN50126, risk assessment should be done at the railway system level. Taking into account that any party that knows the entire complex system is missing, experts explain a challenge when considering the integrated system. For the creation of the total safety architecture, an integrator should create an integral coherence of the safety architecture and the define interdependencies between elements. An overview is necessary to define a complete, comprehensive, and defensible argument

Earlier research explained the need for a comprehensive approach to obtain better understanding of the complex nature of hazards, and understand interrelatedness of all factors that play a role in risk assessment [60]. The complexities of an entire SoS may be more obvious when analysing the overall system architecture and therefore managed more effectively [39]. However, a lack of a system overview can be a major barrier to evolve systems.

12) Risk assessment

Experts indicate knowledge and understanding of the total safety architecture are of primary importance to foresee hazards and risks. The creation of the overall safety architecture requires full knowledge of the risks involved. Organisations manage risk by identification, analysis, and evaluation. Safe operations are achieved by setting and achieving relevant goals.

Without understanding the purpose, goals and decision criteria used to construct and operate systems, it is not possible to completely understand and most effectively prevent accidents [31]. Strong knowledge implies a low degree of uncertainty, and poor knowledge implies a high level of uncertainty [46].

13) Safety

Experts acknowledge complexity in safety. It is assumed that richer understanding of risk assessment improves safety.

Effective management of risk allow an organisation to improve its safety performance. "Practical safety is risk

management" and once that link has been clearly established, the role of safety becomes significant and its value-add more measurable [61].

These described relationships are shown in the upper part of the reference model in Fig. 5. This reference model represents the existing situation on implications with regard to the safety architecture of ERTMS. All ellipses represent key factors, meaning those influencing factors that are considered as the core factors or root causes. The links describe the existing relationships. Not all relationships come from direct experiences. Therefore, these relationships are indicated as an assumption. The existing value of a key factor is represented by means of a '+' or '-'. For example, the '+' next to 'number of different ERTMS inclusions' indicates a high degree. Quantitative classification of whether or not participants experience a challenge in one of the key factors is shown in Fig. 6.

V. DISCUSSION

As for the interviews, we assure the basic quality of the data by forwarding the summary of the interview to each participant, asking for their feedback, and made corrections if there were any misinterpretations.

The reference model is logically verified by consistency, meaning there are no internal conflicts between the key factors, and well-established literature to identify supporting evidence and contradictions. Relationships between safety overview and risk assessment, comprehension of the safety architecture and safety overview, and comprehension and risk assessment, are labelled as assumptions, because they are not described from direct experiences. We verify logically on internal validity and external validity.

As for internal validity of the reference model, causes and effects of a key factor could have been interchanged. For example, a low degree of comprehension of the architecture that leads to a low degree of quality of overview. On the other hand, a low degree of quality of overview that leads to a low degree of comprehension of the architecture. Also, we experienced two causal relationships between fuzzy boundaries and comprehending the architecture, and between various languages and cultures and variety in inclusion ERTMS specs. In the reference model, this is shown through a combined relationship.

As for external validity, findings are person and time-related. Various participants have various experiences. They can interpret a key factor in a different way. Participants with experience between 10 and 20 years will allow fewer generalisation than participants with 5 years of experience and varying backgrounds. Also, statements of participants about topics that fall outside of their expertise are less reliable than the statements of participants that have direct experience with the topic. For these reasons, care is taken that collected data is relevant. Data obtained within the participant's expertise, based on direct experience, is valued higher than the data from an unexperienced participant. We prevent asking leading questions or emphasising a specific detail of the topic. Challenges of key factors and their interrelationships must be suggested by the expert. An

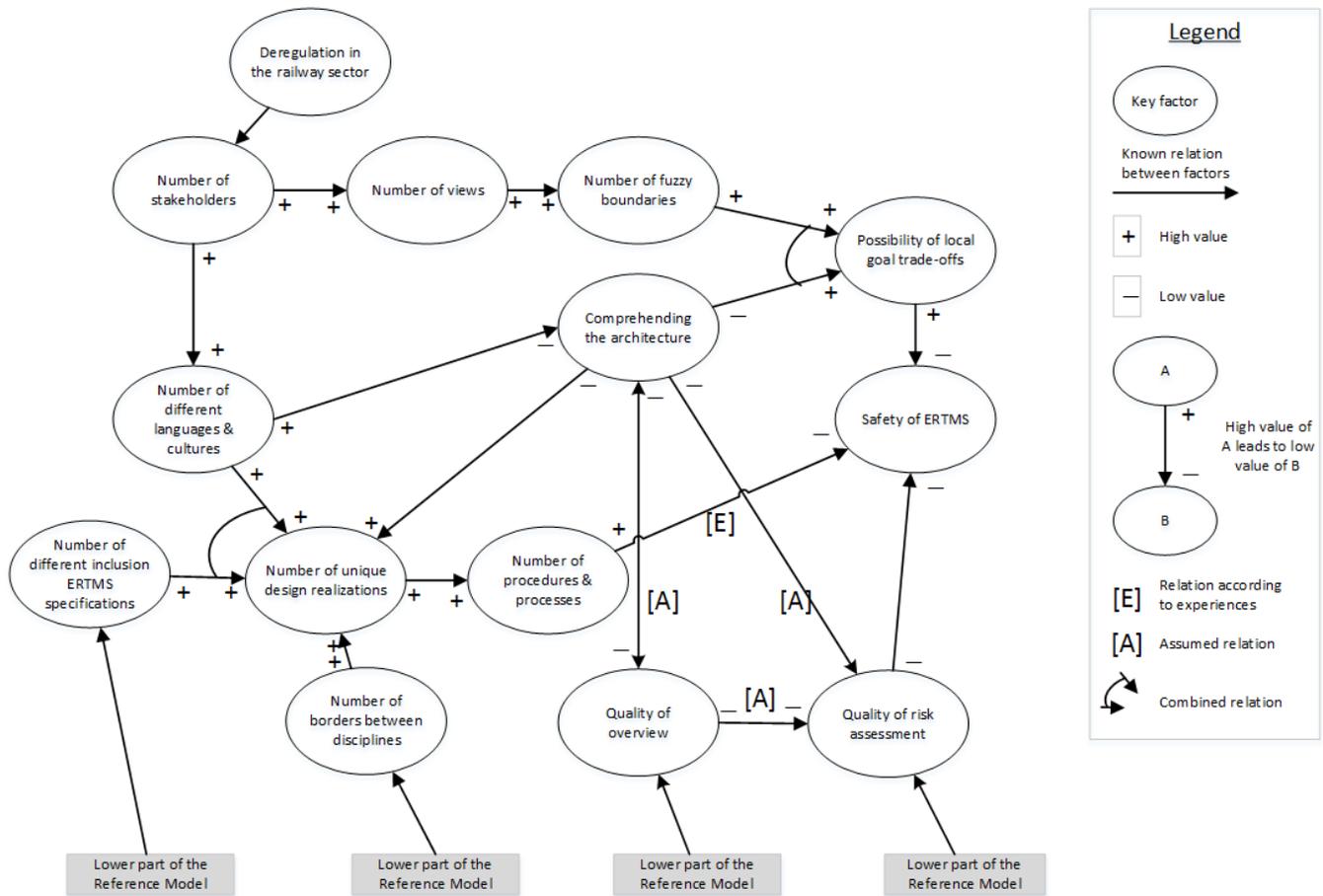


Figure 5. Upper part of the reference model on implications with regard to the safety architecture of ERTMS



Figure 6. Participants experiencing challenges in key factors

explanation of a real-life example or experience suggested by the participant was valued higher.

As for the relationships that are labelled as assumption, we plan to verify them in a real ERTMS case that, at this moment, is setting up their risk analyses and evaluations.

VI. CONCLUSION

Evaluating risks using the ALARP-principle involves complex and challenging explanations and interpretations of the desired performance levels. Time pressure and pressure towards cost-effectiveness can inadvertently lead to generating adaptive responses, wrong/missing identification of hazards and safety risks, and also to safety concessions. Addressing primarily potentially catastrophic risks automatically means accepting the remaining risk.

Both internationally and nationally, stakeholders have different preferences for ERTMS design. In addition, specification interpretations by train operators, infrastructure providers and manufacturers vary markedly.

A system integrator that knows the entire complex system, misses. A system model cannot be built by a bottom-up aggregation of models, derived from research in the individual disciplines, but by a top-down, system-oriented approach based on control theoretic concepts [21]. At the same time, a trade-off using bottom-up aggregation can provide information that helps to focus on more detailed problems of subsystems. Both implicit data-exchange and a missing integral view make it hard to perform a comprehensive safety assessment.

Both international and national preferences, changing specifications, varying stakeholders and various manufacturers led to a unique realisation for every subsystem. Similarly, the occurrence of further transitions with accompanied complexity and procedures and processes that multiply and have wider ramifications. The checking of the critical specifications in natural language is a burdensome task.

Since there is no integral view, local actors can change their conditions without, what at first glance may seem, apparent implications. This in turn leads to a system of a wide variety of subsystems, and the associated increase of transitions that affect safety. In view of this, in order to develop a safe system, the key role is that of the safety architect to define an integral safety architecture, representing:

- Safety functions.
- A top-down risk assessment at the railway system level that relies on system definition, and includes risk analysis and risk evaluation.
- A bottom-up hazard analysis of the system under consideration.

Split-responsibility results in stakeholders that lack insight into cross-border information. Understanding relationships between risk and design can aid in communication between safety engineer and designer. Clear communication regarding safety, which supports critical system development, is essential.

In practice, the unstable specifications and various interpretations are a major problem when dealing with such systems. The consequences are significant: the five ERTMS-projects in the Netherlands (Betuweroute, Port of Rotterdam, HSL South, Amsterdam-Utrecht and Lelystad-Zwolle) are all different [12] in design and use, let alone the wider European variants.

The current integral architecture lacks integrated knowledge, traceability, and consistency. As for safety, this means that the lack of availability of information makes it difficult to find a root cause for each hazard.

Assuming that a safety assessment is conducted in a professional and scientific way, it will meet some standards on quality. For example, all the steps of the assessment are traceable, all assumptions are recorded, and all analysis principles and methods adopted are justified [46]. To meet unforeseen events and surprises, and to identify safety requirements at the system level, a systems engineering approach to safety must be treated adequately in the context of the social and technical system as a whole. As Kecklund [62] also explained: "It became clear that the work of several organisations and authorities at the societal level has implications for railway safety, and therefore, it is important that established channels for communication and co-operation among these parties exist, and that there is a level that affords an overall, holistic perspective."

Current challenges concern an interdisciplinary approach on both the social and technical level, and how parts interact and fit together. Accompanying questions concern the acceptability of the level of incompleteness. ERTMS is a great example of a complex system, subject to an increasing number of stakeholders, various interpretations of requirements, where overall responsibility is split. As for safety, many and varied interactions among the individual components is to be approached proactively and qualitatively where time drain and pressure towards cost-effectiveness can inadvertently lead to generating adaptive responses.

These challenges require improvements in resilience, more awareness and sensitivity for interrelationships between hazards and risks, but even more: comprehending the safety architecture and creating cross-discipline understanding.

In this study, the effects of the safety case regime, interoperability, deregulation and dynamic specifications on the ERTMS have been researched at the Dutch national level.

Achieving an interoperable and safer railway system by implementing ERTMS appears not to be straightforward for three key reasons:

- The safety case argument involves descriptions and observations that require explanations and interpretations from various stakeholders having various points of view, skills, responsibilities and interests into the outcomes of the assessments.
- For the Dutch situation, the absence of central designer and overall processes lowers the degree to which the parties succeed in effectively harmonising various processes.
- An increased number of actors has caused a lack of insight into cross-border information.

Specifications allowing multiple interpretations result in a wide variety of design choices, disparities among systems, possible little recognition of hazards and risks, and needlessly cumbersome procedures.

REFERENCES

- [1] K. Schuitemaker and M. Rajabalinejad, "ERTMS Challenges for a Safe and Interoperable European Railway System," Proceedings of the 7th International Conference on Performance, Safety and Robustness in Complex Systems and Applications (PESARO 2017), pp. 17-22, April 2017.
- [2] European Commission, "ERTMS – European Rail Traffic Management System," Available from: <http://www.ec.europa.eu> [retrieved: March, 2017].
- [3] International Union of Railways, "SafeCulture: A method for assessing organisational safety at interfaces," UIC, 2004.
- [4] R. Forsberg, "Conditions affecting safety on the Swedish railway – Train drivers' experiences and perceptions," Safety Science, vol. 85, pp. 53-59, 2016.
- [5] Ministry of Transport, Public Works and Water Management, "Cost benefit analysis of implementation strategies for ERTMS in the Netherlands," 2010.
- [6] Ministry of Infrastructure and the Environment, "Beleidsimpuls Railveiligheid," 2016.
- [7] Prorail, "Jaarrapport 2015. Analyse STS-passages" Unpublished.
- [8] P. Smith, A. Majumdar, and W. Y. Ochieng, "An overview of lesson learnt from ERTMS implementation in European railways," Journal of Rail Transport Planning & Management, vol. 2, pp. 79-87, 2012.
- [9] F. Laroche and L. Guihéry, "European Rail Traffic Management System (ERTMS): Supporting competition on the European rail network?" Research in Transportation Business & Management, vol. 6, pp. 81-87, 2013.
- [10] M. Ghazel, "Formalizing a subset of ERTMS/ETCS specifications for verification purposes," Transportation Research Part C, vol. 42, pp. 60-75, 2014.
- [11] European Commission, "ERTMS – FAQ on ERTMS," Available from: <http://www.ec.europa.eu> [retrieved: March, 2017].
- [12] Ministry of Infrastructure and the Environment, "ERTMS kennisboek versie 2.0," 2014 Available from: <http://www.rijksverheid.nl> [retrieved: March, 2017].
- [13] Ministry of Infrastructure and the Environment, "Eindrapport Lessen uit het rijden onder ERTMS Level 2 in Dual Signalling omstandigheden," 2015 Available from: <http://www.rijksverheid.nl> [retrieved: March, 2017].
- [14] J. Stoop, J. Vleugel, and J. Baggen, "Testing the untestable: towards pro-active safety assessment," Proceedings of the 9th International probabilistic safety assessment and management conference (PSAM 9), May 2008, pp. 784-791, ISBN: 978-988-99791-5-7.
- [15] L. de Haan, J.L.M. Vrancken, and Z. Lukszo, "Why is intelligent technology alone not an intelligent solution?" Futures, vol. 43, pp. 970-978, 2011.
- [16] OECD. Structural reform in the rail industry. Policy roundtables. 2005.
- [17] G. Alexandersson and S. Hultén, "The Swedish Deregulation Path," Review of Network Economics, vol. 7, issue 1, pp. 1-19, March 2008.
- [18] UNIFE, "UNISIG, An industrial consortium to develop ERTMS/ETCS technical specifications," Available from: <http://www.ertms.net> [retrieved: March, 2017].
- [19] European Union, "Commission Decision of 25 January 2012 on the technical specification for interoperability relating to the control-command and signalling subsystems of the trans-European rail system," Official Journal of the European Union, vol. 55, pp. 1-51, February 2012.
- [20] UNIFE, "A Unique Signalling System For Europe. The long journey to an interoperable railway system," Available from <http://www.ertms.net> [retrieved: March, 2017].
- [21] J. Rasmussen, "Risk management in a dynamic society: a modelling problem," Safety Science, vol. 27, pp. 183-213, 1997.
- [22] N.P. Hoj and W. Kroger, "Risk analyses of transportation on road and railway from a European Perspective," Safety Science, vol. 40, pp. 337-357, 2002.
- [23] B. G. Glaser and A. L. Strauss, "The Discovery of Grounded Theory: Strategies or Qualitative Research," Weidenfeld and Nicolson, London, 1967.
- [24] L. Blessing and A. Chakrabarti, "DRM, a Design Research Methodology," Springer-Verlag London, 2009.
- [25] European Railway Agency, "Intermediate report on the development of railway safety in the European Union," 2013 Available from: <http://www.era.europa.eu> [retrieved: March, 2017].
- [26] K. Schuitemaker, J.G. Braakhuis, and M. Rajabalinejad, "A Model Based Safety Architecture Framework for Dutch High Speed Train Lines," Proceedings of the 10th International Conference on System of Systems Engineering (SoSE), May 2015, pp. 21-29, 2015.
- [27] S. Barua, X. Gao, and M.S. Mannan, "Comparison of prescriptive and performance-based regulatory regimes in the U.S.A and the U.K.," Journal of Loss Prevention in the Process Industries, vol. 44, pp. 764-769, 2016.
- [28] S.Nair, J.L. de la Vara, M. Sabetzadeh, and D. Falessi, "Evidence management for compliance of critical systems with safety standards: a survey on the state of practice," Information and Software Technology, vol. 60, pp. 1-15, 2015.
- [29] ERTMS Strategy Group, "National ERTMS Business Requirements," Available from: <http://www.rssb.co.uk> [retrieved: March, 2017].
- [30] D. Falessi, M. Sabetzadeh, L. Briand, E. Turella, T. Coq, and R. K. Panesar-Walawege, "Planning for Safety Standards Compliance: A Model-Based Tool-supported Approach," IEEE Software, Vol. 29, pp. 54-70, 2011.
- [31] N.G. Leveson, "Engineering a Safer world, Systems Thinking Applied to Safety," The MIT Press, 2011.
- [32] H. Priemus, "Mega-projects: Dealing with Pitfalls," European Planning studies, vol. 18, pp. 1023-1039, 2010.
- [33] T. Pasquale, E. Rasaria, M. Pietro, and O. Antonio, "Hazard Analysis of Complex Distributed Railway Systems," Proceedings of the 22nd International Symposium on Reliable Distributed Systems (SRDS'03), October 2003, pp. 283-292, ISSN: 1060-9857, ISBN: 0-7695-1955-5.
- [34] Tweede Kamer Der Staten-Generaal, "Verslagen van de openbare verhoren," Available from: <http://www.tweedekamer.nl> [retrieved: March, 2017].
- [35] S. Nusser, "Robust Learning in Safety-Related Domains. Machine Learning Methods for Solving Safety-Related Application Problems," Doctoral Thesis, Univ., Fak. Für Informatik, 2009.
- [36] B. Enserink, "Integral assessment – putting safety on the agenda for mitigation and preparedness," Safety Science, vol. 39, pp. 93-105, 2001.
- [37] N.P. Hoj and W. Kröger, "Risk analyses of transportation on road and railway from a European Perspective," Safety Science, vol. 40, pp. 337-357, 2002.
- [38] C. Harvey and N. A. Stanton, "Safety in System-of-Systems: Ten key challenges," Safety Science, vol. 70, pp. 358-366, 2014.

- [39] T. Kelly, "A Systematic Approach to Safety Case Management," SAE Technical Paper, pp. 239-248, 2004.
- [40] N. Yakimets, S. Dhoub, H. Jaber, and A. Lanasse, "Model-Driven Safety Assessment of Robotic Systems," IEEE Intelligent Robots and Systems (IROS), pp. 1137-1142, 2013.
- [41] P. Rook, "Controlling Software Projects," Software Engineering Journal, vol. 1, pp. 7-16, 1986.
- [42] P. Mauborgne, S. Deniaud, E. Levrat, E. Bonjour, J. Micaelli, and D. Loise, "Operational and System Hazard Analysis in a Safe Systems Requirement Engineering Process – Application to automotive industry," Safety Science, vol. 87, pp. 256-268, 2016.
- [43] K. J. Hayhurst and C. M. Holloway, "Challenges in software aspects of aerospace systems," Proceedings of the 26th Annual NASA Goddard, August 2002.
- [44] Lutz, R, "Analyzing software requirements errors in safety-critical, embedded systems," Proceedings of IEEE International Symposium on Requirements Engineering, pp. 126-133, 1993.
- [45] J. Bradley, M. Efatmaneshnik, and M. Rajabinejad, "Toward a Theory of Complexity Escalation and Collapse for System of Systems," Proceedings of the 10th International Conference on System of Systems Engineering (SoSE), pp. 7-11, 2015.
- [46] T. Aven, "Risk, Surprises and Black Swans: Fundamental Ideas and Concepts in Risk Assessment and Risk Management," Routledge, 2014.
- [47] M. Rajabinejad, A. Martinetti, and L. A. M. van Dongen, "Operation, safety and human: Critical factors for the succes of railway transportation," Proceedings of the 11th System of Systems Engineering Conference (SoSE), pp. 1-6, 2016.
- [48] ERA, "Intermediate report on the development of railway safety in the European Union," 2013.
- [49] A. S. Townsend, "Safety Can't Be Measured – An Evidence-based Approach To Improving Risk Reduction," Gower Publishing, 2013.
- [50] J. Iglesias, A. Aranz, M. Cambronero, C. de la Roza, B. Domingo, J. Tamarit, J. Bueno, and C. Arias, "ERTMS deployment in Spain as a real demonstration of interoperability. Near future challenges," 9th World Conference on Railway Research, 2011.
- [51] The Dutch Safety Board, "Train Derailment Hilversum," 2014. Available from: <http://onderzoeksraad.nl/> [retrieved: March 2017].
- [52] A. Laurino, F. Ramella, and P. Beria, "The economic regulation of railway networks: A worldwide survey," Transportation Research Part A, vol. 77, pp. 202-212, 2015.
- [53] I. Sommerville, R. Lock, T. Storer, and J. Dobson, "Deriving Information Requirements from Responsibility Models," Proceedings of the 21st International Conference on Advanced Information Systems Engineering, pp. 515-529, 2009.
- [54] S. Dekker, "Drift into failure: from hunting broken components to understanding complex systems," CRC Press, 2011.
- [55] European Transport Safety Council (ETSC), "Transport accident and incident investigation in the European Union," Available from: <http://www.etsc.eu> [retrieved: March, 2017].
- [56] European Railway Agency, "Analysis of Potential Interoperability Problems Final Report WP4," Available from: <http://www.era.europea.eu> [retrieved: March, 2017].
- [57] European Railway Agency, "Subset-026—2 v300," Available from: <http://www.era.europea.eu> [retrieved: March, 2017].
- [58] G. Baxter and I. Sommerville, "Socio-technical systems: From design methods to systems engineering," Interacting with Computers, vol. 23, pp. 4-17, January 2011.
- [59] N. G. Leveson, "A new accident model for engineering safer systems," Safety Science, vol. 42, pp. 237-270, 2004.
- [60] K. Schuitemaker, M. Rajabinejad, and J.G. Braakhuis, "Model-based safety architecture framework for complex systems," Proceedings of the European Safety and Reliability Conference (ESREL), pp. 3611-3618, 2015.
- [61] S. Merchant, "Role of Safety and Product Integrity," Procedia Computer Science, vol. 8, pp. 443-451, 2012.
- [62] L. Kecklund, E. Olsson, A. Jansson, G. Kecklund, and M. Ingre, "The Train Project: Effects of organizational factors, automatic train control, work hours and environment: suggestions for safety enhancing measures," Proceedings of the Human Factors and Ergonomics Society 47th Meeting, pp. 1835-1839, 2003.

Univariate Modeling of the Timings and Costs of Unknown Future Project Streams: A Case Study

Alireza Shojaei, Ian Flood

M.E. Rinker, Sr. School of Construction Management, University of Florida
Gainesville, Florida, USA.

Email: a.shojaei@ufl.edu, flood@ufl.edu

Abstract—Providing a practical and comprehensive methodology to facilitate management and coordination of multiple projects in a company’s portfolio is a challenging task. Historically, the focus of research has been limited to the selection and prioritization of the set of known projects, current and near future. It is argued that existing portfolio planning models can be improved by adding a stochastic generator of project streams that extends the portfolio and strategic planning horizon to include future unknown projects. The study both identifies the historical factors in the market that are strong predictors of the profile of future project streams and evaluates alternative modeling approaches to the problem. The outputs from the generator are those parameters most critical to a company, namely the occurrence and letting date of a project, its expected duration, and its expected cost. A case study of design-bid-build highway construction projects let by the Florida Department of Transportation (FDOT) is presented for developing, validating and testing the concept of a project stream generator. The results show that FDOT’s future projects can be stochastically forecasted by using historical data and autoregressive moving average modeling along with sampling from representative distributions of cost and durations of FDOT’s projects.

Keywords—Project Portfolio Management; Stochastic Forecasting; Time Series Modeling; Strategic Planning; Uncertainty.

I. INTRODUCTION

Previous work has established the need for future projecting portfolio planning, and proposed some basic modeling approaches to address this issue [1]. This paper advances that work by developing and evaluating the proposed methods of forecasting streams of future work that may be added to a future portfolio.

Construction companies are usually involved in multiple projects at any given time. While different projects progress concurrently, they have different goals and objectives. For instance, some projects may have financial objectives while others may be more focused on marketing or strategic networking. Consequently, a key managerial duty is to allocate resources (such as finances, materials, and personnel) between these concurrently ongoing projects and manage their workflow together to maximize the company’s performance [2]. The process of coordinating multiple projects as such is a challenging task because each incoming project affects all other ongoing projects in terms of their

schedule and progress [3], and without foreseeing these effects, the consequences can be devastating. The goal of this study is to develop a stochastic project stream generator to forecast unknown future projects in order to extend the horizon of strategic planning for construction companies.

The success of a construction company is strongly impacted by its ability to strategically plan for and manage a stream of projects, many of which will overlap in time, and all of which are subject to uncertainty about their occurrence, scope and resource needs. This task can be broadly classified as Project Portfolio Management (PPM). Cooper et al. [4] describe PPM as “...dealing with the coordination and control of multiple projects pursuing the same strategic goals and competing for the same resources, whereby managers prioritize among projects to achieve strategic benefit.” Modern portfolio theory was introduced by Markowitz [5] within the context of finance. McFarlan [6] introduced the concept of PPM in an information technology project management context. He suggested using projects as the elements of a portfolio (instead of investments) to better achieve an organization’s objectives as well as reduce the overall risk that the organization encounters during execution of those projects.

Providing a practical and comprehensive methodology to facilitate management and coordination of multiple projects in a company’s portfolio is a challenging task. There are no appropriate analytical solutions available for dynamic scheduling and resource allocation of project portfolios in real-time [3]. Existing proposed mathematical models (such as those of [7]–[12]) cannot handle the complexity of real world challenges due to a limited consideration of significant uncertainties within their models and a lack of provision for dynamic and real-time analysis. The primary focus of PPM research was initially to improve organizational performance by introducing good practices to choose and prioritize projects and ensure that the right mix of projects was adopted. A recurring theme is the alignment of the projects with the organization’s overall strategy. There is also extensive literature on project selection with a mathematical approach [13]–[16]. In this research, it is not proposed that developed models are incorrect. Instead, it is argued they can be advanced by adding a stochastic project generator to extend the portfolio and strategic planning horizon by forecasting the statistical profile of the stream of unknown future projects. The framework discussed in this paper would allow users to take into account unknown future projects in

their portfolio and strategic planning. From this perspective, it is a novel approach, which by the understanding of the authors has not been done before. Using such an extension would allow the user to plan quantitatively for their portfolio of future projects, as opposed to using a conjecture-based approach.

The rest of this paper is organized as follows. Section II provides a review of the shortcomings of existing PPM models and discusses the impact of uncertainties in PPM. Section III describes the project stream generator and the data used for its development and evaluation. Section IV discusses the modeling approach and results. Section V presents the conclusions and identifies future directions for the research.

II. PROJECT PORTFOLIO MANAGEMENT AND UNCERTAINTIES IN STRATEGIC PLANNING

Selecting projects from available options and planning and scheduling for them collectively have recently received a considerable amount of attention [17]. For construction related organizations, such as investors, developers, and contractors, it is critical to gather and analyze project information to select the best options according to their strategic goals and schedule them within the required timeframe and financial constraints. This is a complex and multifaceted process, which has many contributing factors, such as the market condition, the organization's structure, resource availability and so on [18]. Research on this topic has come from several different points of view, including selection model criteria and scheduling mechanisms [19], yet the primary focus has been choosing the most appropriate projects rather than providing a real-time dynamic model to address the project selection and scheduling issues [3]. Another shortcoming has been to disregard the importance of multiple project scheduling and resource allocation under influential factors and uncertainties, such as the economic situation of the construction industry and companies' organizational changes. Despite the wide range of available modeling approaches, companies still struggle to optimize and manage changes among their projects [19]. One of the reasons for this is that the proposed mathematical models cannot address the complexity of the real world situation [3]. Excluding uncertainties (such as the impact of possible upcoming projects) or changes in the economic and financial situation of the construction industry are some other noteworthy contributing factors to the poor performance of existing models.

The concept of uncertainty is very significant within the field of project portfolio management. This has led to an extensive literature on uncertainty and the ways to manage it. Duncan [20] and Daft [21] demonstrated that changes in the business environment combined with projects with high complexity always result in an increase in uncertainty in parameters, such as the number of projects, their performance, and their adherence to the project plan. Farshchian and Heravi [22] used agent-based modeling to evaluate time and cost uncertainties related to current projects on a project portfolio.

The impact of uncertainty on organizations is well established across many disciplines from psychology to economics [23]. Environmental uncertainties and their relation to organizations are analogous to the state of a person with a shortage of critical information about the environment. Scott [18] provides an example of the definition of environmental uncertainty as variability or the extent of predictability of the environment where work is executed. They also introduce some measures for uncertainty, such as variability of inputs, the number of deviations in the work process, and the number of changes in the main products. In the project management context, uncertainty in a project is defined as the accuracy of predicting the variation of resource consumption, output, and work process. Uncertainty in a project can be seen as a variation from expected performance of the system under investigation.

The Project Management Institute (PMI) standard for portfolio management despite introducing the risk management concept at a portfolio level does not provide much information on how managers should handle uncertainty and risk within their portfolio. They only provide guidelines on categorizing different possible stages and processes plus naming some of the possible techniques available to handle uncertainties. The PMI only suggests monitoring risks and the performance of the project portfolio under the monitoring and control process group. The proposed framework by the PMI also includes monitoring changes in business strategy. This is an important task because when it occurs, it might result in a complete realignment of the portfolio. The mechanisms involved in this realignment are not specified other than restarting the whole PPM process from the beginning. Also, ad-hoc disturbances to the ongoing and approved project portfolios are almost entirely neglected. This oversight is not because the topic lacks interest or that authors assume a stable and predictable environment. Rather, it can probably be explained by the fact that the subject of PPM is relatively young and that the researchers and academics preferred to focus on more pressing issues in this area. For many companies, the environment is unstable, and the high level of uncertainty and unknowns resulting from the dynamic environment lead to some challenges. Upcoming projects significantly affect the performance of a project portfolio [3]. The typical approach when a new project is added to the portfolio is to update the project portfolio's plans and to try to re-optimize everything.

III. UNKNOWN FUTURE PROJECT STREAM GENERATOR

This paper presents an approach to statistically represent unknown future projects to extend the portfolio and strategic planning horizon. Forecasting a company's unknown future projects can be based on the company's past and current portfolio data, or it can use historical data from market to forecast all the upcoming projects as project streams and filter those by bidding success models. In an environment, where the supply of the projects is scarce and very competitive, using just the company's past projects to forecast the future unknown projects is potentially less

accurate. Arguably it is more valid to forecast streams of unknown projects (all the available projects in the future) considering the uncertainties in the context and filter those projects by bidding success models to get the final future projects in a company’s portfolio. The forecast can statistically generate a single set of outputs or stochastically produce streams of values as output. Considering the uncertainties in the market, the PPM context, and the availability of future projects, stochastic forecasting appears to be the right choice.

This paper reports on the development, validation, and testing of a project stream generator for design-bid-build highway construction projects let by the Florida Department of Transportation (FDOT). The primary data for this study were obtained from FDOT’s historical project lettings database covering 14 years (from 2003 to 2017). The last two years (2015 and 2016) data are withheld to be used as a validation set for the final model and not being used in this study to be used after more models are tested for final verification without any kind of bias. Thus, the model training and selection are based on the data from 2003 to 2015, which contains 2,816 design-bid-build project-letting reports. The outputs from the generator are those parameters most critical to a company, namely the occurrence and letting date of a project, its expected duration, and its expected cost. Other factors, such as economic condition can have an impact on the project stream. Table I shows a pool of candidate variables containing 24 potentially relevant predictors including the macroeconomics metrics and construction indices that were compiled from the related

sources and literature [24]–[26] that can be used in multivariate modeling.

The data should be split into three sections as a training set, a test and model selection set, and a final validation set for the final model. In this process, different models are trained and tested using the cross-validation method and the best model is validated with the withheld data. The final validation set is the data from 2015 and 2016, and the data from 2003 to 2015 is used for training and testing of different models to find the best performing model and optimize its corresponding values.

The data under study is a time series and so the continuity of the data is important and should not be tampered with by randomly dividing into different sections for validation. As a result, a rolling forecast origin and a rolling window method is used to cross validate the models’ performance to avoid overfitting and overestimation of performance. The rolling window method has a fixed window (Figure 1-A), where the training (orange bar) and test (blue bar) sets duration is fixed and rolls through time. In this research, the training and testing set were chosen to be three years each and roll one year in each trial. The rolling forecast method on the contrary uses progressive length of data as the training set (as shown in Figure 1-B) in each trial. The initial training set was chosen to be three years and increase one year in each trial while the test set remain three years of consecutive data after the end of the training set for each trial. Using both methods can help better understand the model’s performance and give more insights into the characteristics of different time spans of the data.

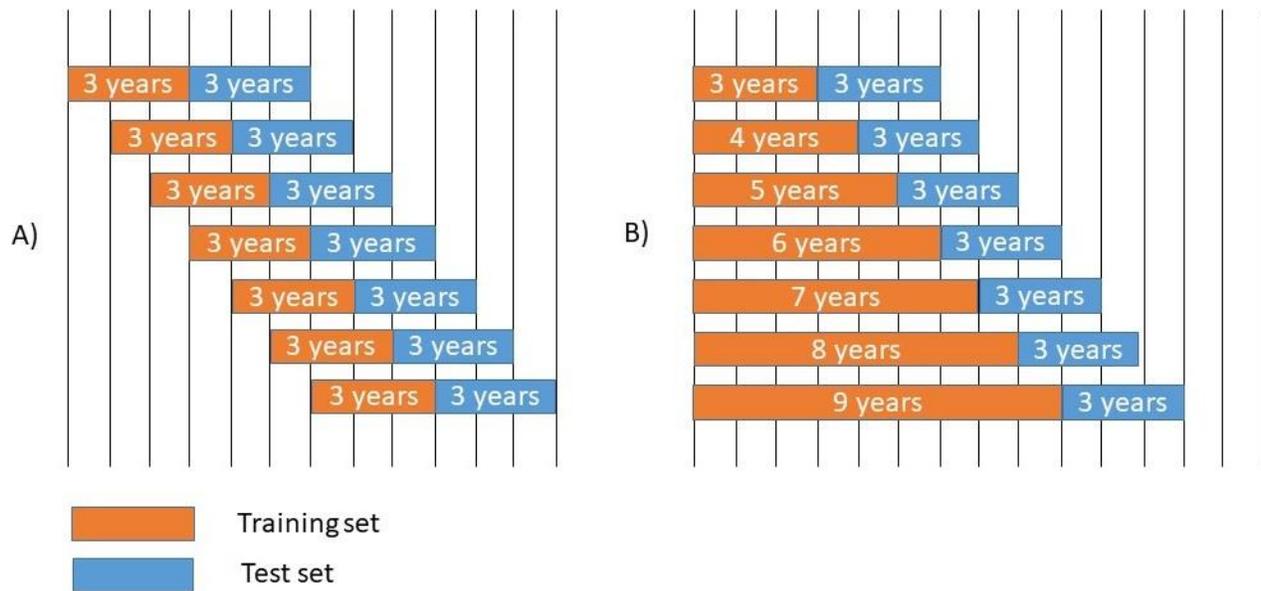


Figure 1. Visual representation of cross-validation methods used. A) Evaluation based on a fixed window rolling forecast B) Evaluation based on an increasing window rolling forecast

TABLE I. Potentially Relevant Predictors.

CANDIDATE VARIABLES	SOURCE
GROSS DOMESTIC PRODUCTS (GDP)	U.S. Bureau of Economic Analysis
GDP IMPLICIT PRICE DEFLATOR	U.S. Bureau of Economic Analysis
INFLATION RATE	World Bank
CONSUMER PRICE INDEX	U.S. Bureau of Labor Statistics
NATIONAL HIGHWAY COST INDEX (NHCCI)	U.S. Department of Transportation
FDOT'S ANNUAL BUDGET	Florida Department of Transportation
FDOT'S PRODUCT BUDGET	Florida Department of Transportation
FEDERAL FUNDS RATE	Federal Reserve Systems
UNEMPLOYMENT RATE	U.S. Bureau of Labor Statistics
FLORIDA UNEMPLOYMENT RATE	U.S. Bureau of Labor Statistics
NUMBER OF EMPLOYEES IN CONSTRUCTION	U.S. Bureau of Labor Statistics
NUMBER OF EMPLOYEES IN CONSTRUCTION IN FL	U.S. Bureau of Labor Statistics
AVERAGE WEEKLY HOURS	U.S. Bureau of Labor Statistics
PRIME LOAN RATE	Federal Reserve System
BUILDING PERMITS	U.S. Bureau of Census
MONEY SUPPLY	Federal Reserve System
AVERAGE HOURLY EARNINGS	U.S. Bureau of Labor Statistics
EMPLOYMENT COST INDEX (ECI) CIVILIAN	U.S. Bureau of Labor Statistics
DOW JONES INDUSTRIAL AVERAGE	Yahoo Finance
CRUDE OIL PRICE	U.S. Energy Information Administration
BRENT OIL PRICE	U.S. Energy Information Administration
PRODUCER PRICE INDEX	U.S. Bureau of Labor Statistics
HOUSINGS STARTS	U.S. Bureau of Census

The sequence of generating information in the proposed model starts with forecasting the number of projects (project frequency) for the desired time span, using the optimal model based on the training and validation from historical data. This is followed by sampling the project costs from the cost distribution. At each point in time, the number of samples from the distribution is based on the number of projects forecasted in the previous step. Finally, the project durations are sampled from the duration distribution. One important issue is the relationship between these main variables in this process. No logical relationship can be established between duration of the projects and the frequency of the projects. However, the frequency of the projects and their accumulated cost has a high correlation, which can be used in the modeling process. Figure 2 represents the four possible ways that this relationship can be accounted for. One option is using a unidirectional assumption to use cost as an exogenous variable to forecast frequency (number one) along with other variables or vice versa (number two). The third option is to use a recursive model and test for convergence of the values. The last option is to neglect this correlation and assume that it is captured through the individual forecasting of each variable. Another important relationship is the possible correlation between cost and duration, which should be considered in the sampling process from their representing distributions. This could be done by using an empirical copula to build a multivariate probability distribution. As a result, the two variables are assigned simultaneously in each round of sampling with the correlation incorporated in the values.

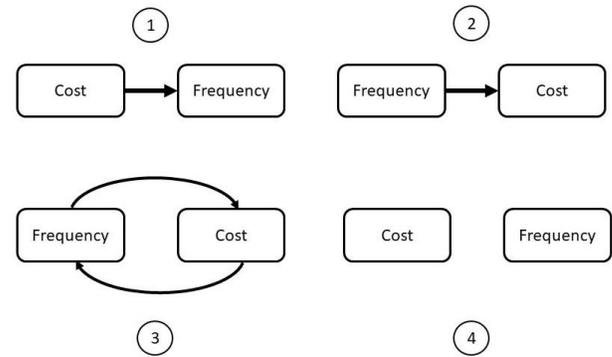


Figure 2. Possible strategies to address the relationship between cost and frequency of the projects. 1) Use cost to forecast frequency 2) Use frequency to forecast cost 3) Recursively use cost and frequency to forecast each other 4) Ignoring the relationship of the cost and frequency in the model

In general, the generator could be implemented as a univariate or multivariate model, and with linear or nonlinear relationships between the inputs and outputs of each model. The complete set of results from the proposed framework can be used as a component in any PPM model to consider unknown future projects in strategic planning.

IV. MODELING APPROACH

Different approaches have been used for time series modeling. Cargnoni et al. [27] used Gaussian models to forecast the number of high-school students in each grade in future school years in the Italian school system. Voyant et al. [28] employed a multilayer perceptron to forecast global solar radiation. Li and Chen [29] used a LASSO (Least Absolute Shrinkage and Selection Operator) based regression to estimate macroeconomic time series, and they demonstrated how this method could be combined with a dynamic factor model to yield a more accurate forecast performance. Exterkate et al. [30] used kernel ridge regression as a multivariate model for economic time-series forecasting by considering the nonlinear relationships among the variables. They found that this method outperformed traditional time-series forecasting techniques based on principal components. Yu and Liang [31] compared the linear ridge regression, ARIMA (Autoregressive Integrated Moving Average), naïve, inverse approach, and support vector machine in forecasting hydrologic time series and concluded that the ridge linear regression outperformed the other models in terms of both performance and time of execution. Choubin et al. [32] compared multiple linear regression, a multilayer perceptron neural network, and an adaptive neuro-fuzzy inference system for forecasting precipitation and concluded that the multilayer perceptron neural network outperformed the other methods. Cao and Tay [33] used a support vector machine for financial time-series forecasting and compared it with a multilayer back-propagation neural network and a regularized radial basis function neural network. They concluded that the support vector machine outperformed the back-propagation neural network and produced a performance similar to that of the regularized radial basis function neural network. The review

of literature shows that dependent on the problem and data, different models perform better. As a result, a set of different models using a systematic approached should be tested to make sure an appropriate model is used for the final forecast.

The scheme used to develop the model is shown in Figure 3. The purpose of this scheme is to look for characteristics of data, to capture them in the model's projections, and then to check to see if the model reproduces them by using the cross-validation tests discussed above. The univariate model, being the simplest, was adopted as a benchmark against which the more complex multivariate models could be compared in terms of forecast accuracy.

The first step is modeling the main variables through univariate modeling methods, such as Autoregressive (AR), Moving Averages (MA), Autoregressive Moving Average (ARMA), and exponential smoothing. More sophisticated approaches such as artificial neural networks can also be implemented considering the availability of the necessary data size to properly the train neural network. After establishing a benchmark, potentially relevant predictors were identified to populate a pool of candidate independent variables based on a literature review and cognitive theories. This introduces the environmental uncertainties to the forecast with the aim of improving the accuracy of the simulation. These variables are not going to have necessarily a causal relationship with the main variables; the only concern here is to be helpful in forecasting the dependent variable.

The next step is exploratory data analysis. It starts with a graphical comparison of the independent and dependent variables, such as scatterplots of pairs of variables. Pearson correlation, unit root (stationary or non-stationary test), Granger causality (helpful for short term forecasting), and cointegration (helpful for long term forecasting) tests are among diagnosis techniques that are relevant.

The last step is to choose a set of multivariate modeling approaches based on the result of the exploratory data analysis and test whether including explanatory variables and models that are more complex can improve the accuracy of the forecast. The range of the models should test for linear and non-linear relationships based on the result of the previous step along with variable selection (pruning), parameter optimization and finding the appropriate lag between variables.

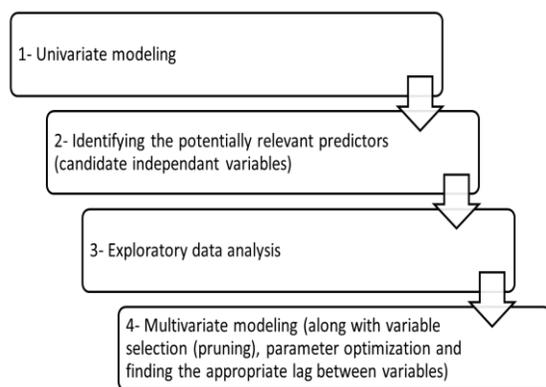


Figure 3. Model development scheme.

Models concerning time series data frequently involve using the value from one or more previous time steps to forecast values at succeeding points in time; in other words, they regress based on past values. In conventional modeling, the assumption is that the independent values are known, and the dependent values are forecast. However, in multivariate time series forecasting, even the independent variables' values in the future are unknown and need to be forecast. As a result, the model contains a system of equations that forecast both independent and dependent variables in the future. This system is recursive when all the causal relationships are unidirectional and non-recursive (simultaneous) when there is reciprocal causation between variables.

Figure 4 shows four of the possible internal structures of the model. Figure 4-A shows the dependencies between the inputs and output in a univariate AR model with a lag of two. In this example, the forecast value at each point in time is based on the two preceding past values. Equation (1) shows the mathematical relationship in such a model, where each value in time is calculated with a linear combination of the past two values plus a constant term (β_0) and a white noise term (ϵ_t).

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_t \quad (1)$$

Figure 4-B shows a recursive multivariate model, where the dependent variable forecast is based on past values of itself and the independent variables. However, each independent variable is only based on its past values. Equation set (2) shows the relationship of such a model with only one lag. It should be noted that in practice there can be many more lags involved in this model and the next two models.

$$\begin{aligned} Y_t &= \beta_{10} + \beta_{11} Y_{t-1} + \beta_{12} X_{t-1} + \beta_{13} X'_{t-1} + \epsilon_{t,1} \\ X_t &= \beta_{20} + \beta_{21} X_{t-1} + \epsilon_{t,2} \\ X'_t &= \beta_{30} + \beta_{31} X'_{t-1} + \epsilon_{t,3} \end{aligned} \quad (2)$$

Figure 4-C shows another recursive model, which differs from model 4-B in that the independent variables also act as input to each other.

$$\begin{aligned} Y_t &= \beta_{10} + \beta_{11} Y_{t-1} + \beta_{12} X_{t-1} + \beta_{13} X'_{t-1} + \epsilon_{t,1} \\ X_t &= \beta_{20} + \beta_{21} X_{t-1} + \beta_{22} X'_{t-1} + \epsilon_{t,2} \\ X'_t &= \beta_{30} + \beta_{31} X_{t-1} + \beta_{32} X'_{t-1} + \epsilon_{t,3} \end{aligned} \quad (3)$$

Figure 4-D shows a sample of a non-recursive (simultaneous) model, where all the variables work as inputs for each other. There is no discrimination between dependent and independent variables in this approach. In this case each variable is a function of its past values and other variables past values.

$$\begin{aligned} Y_t &= \beta_{10} + \beta_{11} Y_{t-1} + \beta_{12} X_{t-1} + \beta_{13} X'_{t-1} + \epsilon_{t,1} \\ X_t &= \beta_{20} + \beta_{21} Y_{t-1} + \beta_{22} X_{t-1} + \beta_{23} X'_{t-1} + \epsilon_{t,2} \\ X'_t &= \beta_{30} + \beta_{31} Y_{t-1} + \beta_{32} X_{t-1} + \beta_{33} X'_{t-1} + \epsilon_{t,3} \end{aligned} \quad (4)$$

Figure 5 shows a summary of the possible univariate models considered in this study. At a high level in this classification are two options, neural networks and time series modeling, each with its own set of variations. A long-short term memory is used as the neural network model and AR, MA, ARMA, and different smoothing methods are used as the time series methods.

Figure 6 shows the possible multivariate models that can be used in similar studies. The top categories here are regression, neural networks, time series, and nonlinear autoregressive moving average with exogenous variables (NARMAX), which is a combination of neural networks and time series models.

After training and validating different models, some diagnostic tests should be conducted to check the stability of the best performing model before its implementation. For instance, checking to see if there is an autocorrelation between the residuals of the forecast is an appropriate tool for time series forecasts. Also, checking the way error compounds and undertaking a sensitivity analysis to see how the values of model parameters affect the model's output can give a deeper insight into the performance of the model.

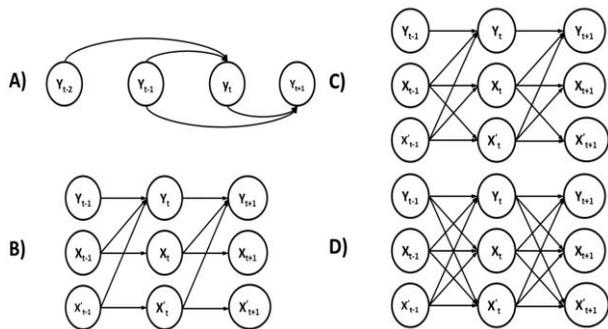


Figure 4. Possible internal structures of the model, illustrating the relationship between the dependent and independent variables. A) Sample illustration of a univariate model B) Sample illustration of recursive multivariate model type 1 C) Sample illustration of recursive multivariate model type 2 D) Sample illustration of recursive multivariate model type 3

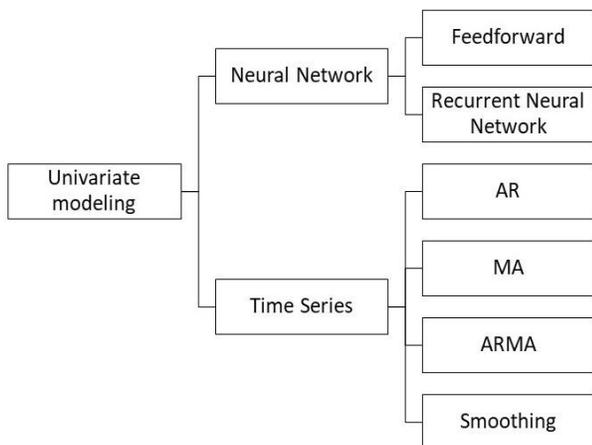


Figure 5. Univariate models surveyed through this study.

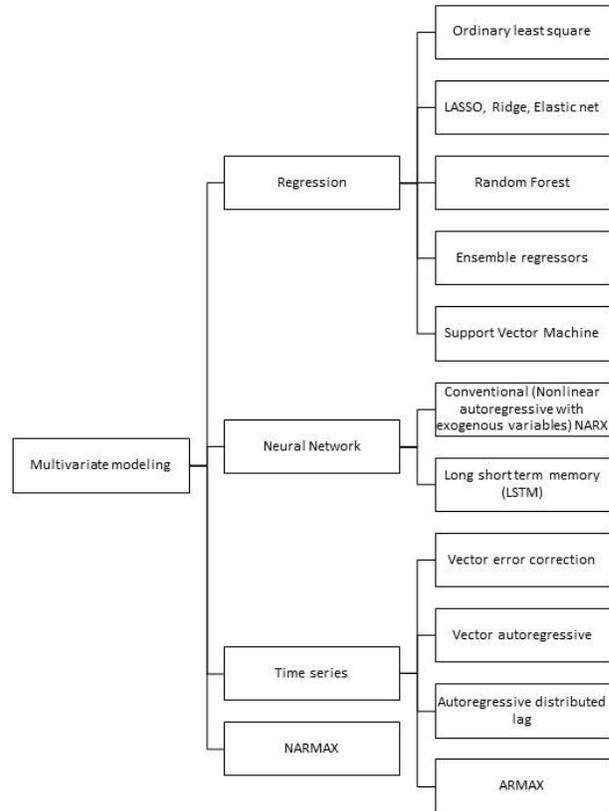


Figure 6. Multivariate models surveyed through this study.

A. Modeling Project Frequency

Before modeling the project frequency, it is necessary to conduct some preliminary data analysis to quantify the data's characteristics. Correlogram of autocorrelation and partial autocorrelation reveals that lag 8 and 12 exceeds the significance bounds, which means extending past 8 and 12 values in univariate modeling are the most appropriate options as they demonstrate significant correlations with the original time series under study.

Testing the stationarity of the project frequency is also important. Figure 7 shows the rolling mean and standard deviation of project frequency plotted along with the actual data. It is visually plausible that the data fluctuate around a fixed mean and variance. It can be numerically assessed by using an Augmented Dickey-Fuller test (ADF) to see if the data is stationary. There are three variations of the ADF test, all with the null hypothesis that a unit root is present in a time series sample (series is not stationary). If under any of the three variations the null hypothesis is rejected it can be inferred that the time series is stationary. The ADF test's result (the appropriate lag is chosen based on the Akaike Information Criterion (AIC)) shows that the null hypothesis can be rejected at a 95 percent confidence level. Therefore, the frequency series is considered stationary, meaning it is evolving around a constant mean and variance.

Two approaches can be implemented to forecast project frequency: univariate and multivariate modeling. ARMA and

exponential smoothing are among the most widely used methods to model a univariate time series. ARMA is used to model stationary time series data and is typically represented as ARMA (p,q), where p is the autoregressive order and q is the moving average order. The order of autoregressive and moving average is selected via autocorrelation and partial autocorrelation correlograms. Based on the preliminary data analysis of project frequency an ARMA (p=8, q=8) is the best choice to model the project frequency series. Also, a set of seasonal ARMA models fitted to the data and the best model is selected via AIC. Moreover, other univariate time series methods such as AR, MA, exponential MA, double exponential MA, different variations of ARMA, triple exponential smoothing (Holt-Winters, which takes into account both seasonal changes and trends) method are implemented. This analysis is conducted on the trained and validation data set from 2003 to 2015 using both rolling origin and rolling window methods.

The performance on the test set is the critical measure to compare the performance of the models. Table II presents the summary of the best univariate models and their performance to forecast the project frequency measured by both rolling origin and rolling window cross-validation methods using Root Mean Squared Error (RMSE). Table III presents Mean Absolute Error (MAE) of the same models and cross-validation methods. It should be noted that the results represented here are the average error of the trained models tested on the seven test data sets presented in the cross validation methods of Figure 1. Each test set consist of three different years, so it is safe to assume that there is no over-parametrization or over-fitting represented in the average errors. The results show that almost all the models perform better according to the rolling origin method. This could be due to the fact that in this method the training data is more than what is being used in the rolling window. Thus, the coefficients are calculated more appropriately. Comparing the results of the rolling origin cross-validation method across different models shows that the ARMA (8,8) model outperformed the other models for both the RMSE and MAE measures.

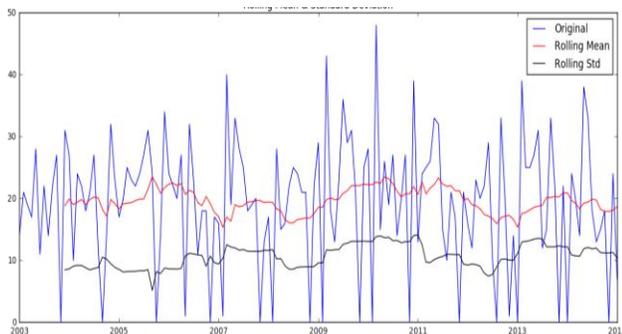


Figure 7. Rolling mean and standard deviation of project frequency plotted against the original data.

TABLE II. SUMMARY OF RMSE OF TIME SERIES MODELS.

Model	Average (rolling origin)	Average (Rolling window)
AR (8)	10.925	10.993
AR (12)	10.934	11.063
MA (8)	11.321	11.343
MA (12)	11.288	11.308
Exponential MA (8)	11.404	11.420
Exponential MA (12)	11.324	11.325
Double Exponential MA (8)	12.050	12.057
Double Exponential MA (12)	11.647	11.683
Auto ARMA	11.057	11.127
ARMA (8,8)	10.715	11.580
ARMA (8,12)	10.830	11.252
ARMA (12,8)	10.870	12.003
ARMA (12,12)	11.556	42.616
Exponential smoothing	11.057	11.138
Holt Winter	10.820	12.814

TABLE III. SUMMARY OF MAE OF TIME SERIES MODELS.

Model	Average (rolling origin)	Average (Rolling window)
AR(8)	8.48	8.551
AR (12)	8.49	8.533
MA (8)	8.85	8.857
MA (12)	8.74	8.765
Exponential MA (8)	9.02	9.040
Exponential MA (12)	8.89	8.894
Double Exponential MA (8)	9.75	9.693
Double Exponential MA (12)	9.3	9.320
Auto ARMA	8.59	8.611
ARMA (8,8)	8.45	9.297
ARMA (8,12)	8.53	9.218
ARMA (12,8)	8.55	9.739
ARMA (12,12)	9.23	31.853
Exponential smoothing	8.59	8.611
Holt Winter	8.7	10.158

The evaluated models so far presented focus only on linear relationships between the inputs. In order to investigate the nonlinear relationship between the inputs a Long-Short Term Memory (LSTM) neural network is used. This method is only implemented with the rolling origin cross-validation method as it requires more data for comprehensive training and the data limitation of the rolling window would reduce its performance significantly. Table IV shows the summary of the different LSTM models trained and tested using a grid search to find the optimal number of neurons and lookback number (number of previous values to be considered as input). The results show that the net with two neurons and one lookback performs better than other configurations.

TABLE IV. SUMMARY OF LSTM MODELS PERFORMANCE USING ROLLING ORIGIN METHOD.

lookback	neurons	Average (RMSE)	Average (MAE)
1	1	10.79	8.56
1	2	10.75	8.58
1	3	10.76	8.61
1	4	10.76	8.61
1	5	10.79	8.63
1	10	10.76	8.59
1	20	10.77	8.61
3	1	11.18	8.81
3	2	11.52	9.15
3	3	11.22	9.12
3	4	11.42	9.12
3	5	11.33	9.22
3	10	12.01	9.65
3	20	12.36	9.70
5	1	11.81	9.37
5	2	11.54	9.05
5	3	11.36	9.09
5	4	11.66	9.26
5	5	11.27	9.01
5	10	12.78	10.28
5	20	14.25	11.00
8	1	12.43	10.01
8	2	13.21	10.74
8	3	15.80	12.45
8	4	15.81	12.70
8	5	16.78	13.25
8	10	18.71	14.55
8	20	21.27	15.85
12	1	17.75	13.02
12	2	17.78	13.51
12	3	17.81	13.99
12	4	19.54	15.41
12	5	18.19	14.85
12	10	15.98	13.16
12	20	16.94	13.40

The difference between the performance of the models might seem small. However, considering that it is an average of seven test sections based on the cross validation methods discussed earlier, even small differences are meaningful. Comparing results of the time series model and the LSTM model shows that ARMA (8,8) is the best approach for modeling project frequency. Table V presents a quantitative summary of the training and test set of the last cross validation section for a better understating of how the model

compares to the actual data. The mean and median match very well for both the training and test sets. However, the model's variance, standard deviation and range are less than the actual data. Figure 8 provides a more in-depth understanding of the results by a visual illustration of the performance of the ARMA model, illustrating the difference between the actual data and the best performing model. The predicted values are shown in blue, and the actual data are plotted in red. Visual inspection of Figure 8 shows that the model performs better forecasting later values (after 2008) and, likewise, better captures the variance of the actual data in these later years. However, it is evident that the model's variance (blue) is less than the actual data (red) through the whole data set. The gray area represents the prediction intervals for the test data set. The dark grey shows the 80% interval and light grey shows the 95% interval.

Based on the literature [24], [34], [35] including explanatory variables and using multivariate models can yield more accurate results. As a result, following the scheme illustrated in Figure 3 (using multivariate methods to improve project frequency forecast) is part of future work in this study.

B. Modeling Cost and Duration

Cost and duration are the two variables to be sampled from a fitted distribution from past projects. Checking for the correlation between the two variables is essential. A Pearson correlation test shows 0.662 correlation coefficient with 0.00 P-value between the duration and cost at the project level (0.00 P-value shows that the correlation is significant, and it is not due to the chance). This shows a moderately linear relationship between the two variables, and it should be incorporated in the model.

Each member in a set of continuous distributions (consisting of the Inverse Gaussian, Pearson, Fréchet, Normal, Lognormal, Dagum, Fatigue Life, Logistic, Loglogistic, Gamma, Exponential, Triangular, Uniform, Student, and Weibull distributions) has been parametrically fitted to the cross-validation data sections using the maximum likelihood estimation (MLE) method and ranked via AIC. Table VI shows the results from the rolling window cross-validation while Table VII shows the result for the rolling origin cross-validation method. From these tables it can be seen that the distributions that were consistently among the best were the Inverse Gaussian distribution for duration, and the Lognormal distribution for cost.

Figure 9 shows the histogram, and the corresponding fitted distribution for the duration and cost of the projects. An Inverse Gaussian distribution with $\mu=244.67$ and $\lambda=273.93$ was found to provide the best fit using AIC for the duration. A lognormal distribution with (mean log) $\mu=14.413319$ and (standard deviation log) $\sigma=1.524961$ was found to provide the best fit using AIC for the cost. Through sampling from these distributions, a cost and a duration can be assigned to each forecasted project. As a result, the output of the framework would be the number of projects for each month and a cost and a duration assigned to each project.

TABLE V. QUANTITATIVE SUMMARY OF ARMA MODEL AND ACTUAL DATA

	Training set		Test set	
	ARMA model	Actual data	ARMA model	Actual data
Mean	19.83	19.82	19.92	18.56
Variance	31.33	104.97	38.89	121.57
Std. Dev.	5.60	10.25	6.24	11.03
Median	20.58	21.00	20.96	20.00
Minimum	1.09	0.00	6.87	0.00
Maximum	30.78	48.00	30.75	39.00
Range	29.69	48.00	23.87	39.00

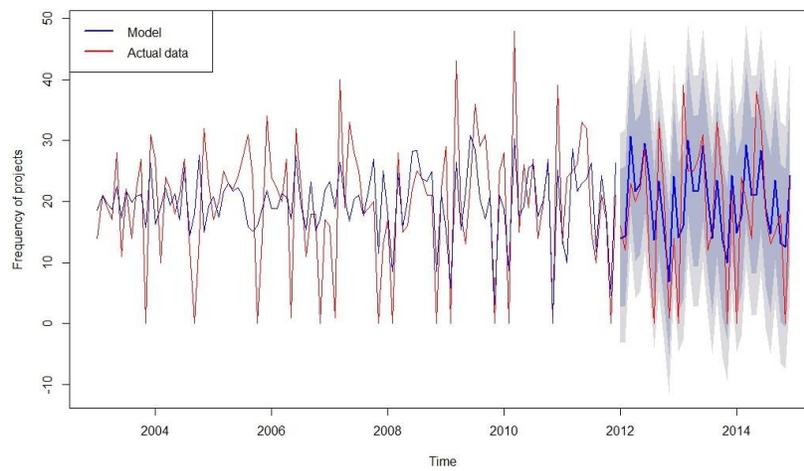


Figure 8. ARIMA (8,0,8) model illustration using 9 years for training and 3 years for testing.

TABLE VI. DISTRIBUTION FIT RESULTS USING ROLLING WINDOW DATA SECTIONS

		2003-2008	2004-2009	2005-2010	2006-2011	2007-2012	2008-2013	2009-2014
Cost	Best distribution	Lognormal						
	AIC	45,368.30	46,444.59	47,136.99	46,065.73	45,450.35	45,732.72	45,863.51
Duration	Best distribution	Inverse Gaussian						
	AIC	17,724.74	18,115.82	18,322.96	17,963.08	17,821.47	17,933.65	18,022.70

TABLE VII. DISTRIBUTION FIT RESULTS USING ROLLING ORIGIN DATA SECTIONS

		2003-2008	2003-2009	2003-2010	2003-2011	2003-2012	2003-2013	2003-2014
Cost	Best distribution	Lognormal						
	AIC	45,368.30	53,914.56	62,259.21	69,807.80	76,323.25	84,308.77	91,330.68
Duration	Best distribution	Inverse Gaussian						
	AIC	17,724.74	21,099.32	24,383.72	27,346.81	29,928.68	33,040.76	35,774.54

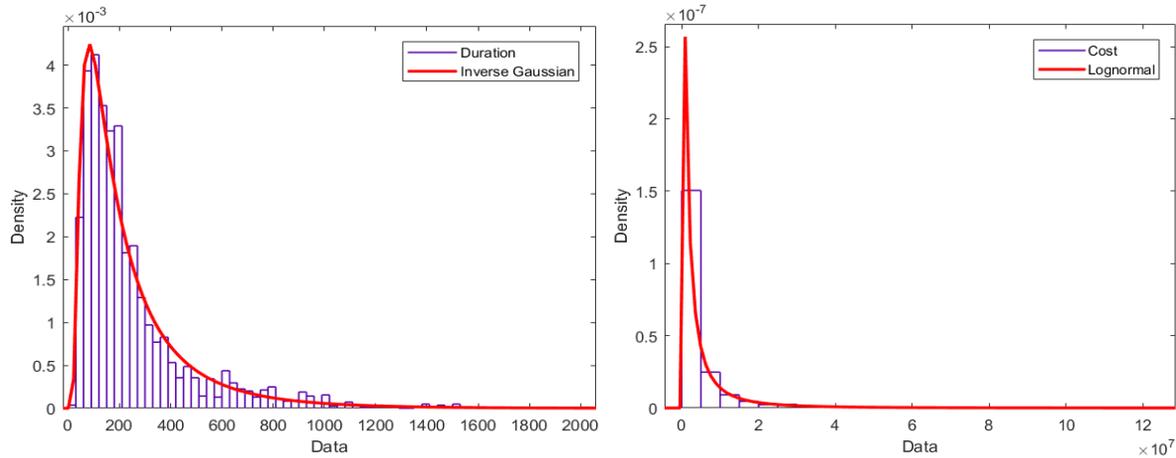


Figure 9. Duration (Left) and Cost (Right) fitted distributions

The performances of the various model components presented in this section indicate the viability of an integrated project stream forecaster that predicts, within a simulation environment, the frequencies of projects and empirical distributions of project duration and cost. Specifically, the generator will produce stochastic streams of unknown future FDOT projects.

V. CONCLUSION AND FUTURE WORK.

This paper has proposed an extension to the body of existing project portfolio planning models and discussed a methodology for its development. The proposed model will extend the horizon for portfolio and strategic planning by enabling users to look further into the future and consider unknown (but statistically quantifiable) projects alongside the known and current projects in their planning process.

The proposed model provides an additional component to the current portfolio management models. A general modeling approach with different possible training and validating methods is discussed and results of the research on developing, validating and testing a stream generator to forecast FDOT projects, in terms of time of occurrence, expected duration and expected cost, is presented. It is shown how univariate models can be used to forecast project frequency, and a discussion is provided of the representing distributions for project cost and duration along with their relationship. Results of project frequency univariate modeling showed that ARMA was the best performing model in this case, outperforming the LSTM neural network. This could be explained by the excessive need of such neural networks for large sample datasets. Furthermore, among the tested distributions, the Inverse Gaussian was found to be the most representative of project duration, and the Lognormal distribution was found to be the most representative distribution for project costs.

A set of potentially relevant predictors including the macroeconomics metrics and construction indices have been identified to enable future improvement of the model

using multivariate methods. This approach can be applied in different contexts and is not confined to the specific case study discussed in this paper. It is also proposed to expand the scope of the research by adding other characteristics to the project stream generator (such as different project types) and implementing it within various environmental contexts.

The complete framework will allow the user to examine different bidding and project selection strategies to see the impact on a company's portfolio and the future resource demands. Furthermore, it will lead to the selection of a closer to optimal strategy and optimal resource distribution for a user. Finally, taking into account uncertainties in future project streams might decrease the required extent of continuous adjustments to a company's portfolio plan resulting from new projects being added to the portfolio.

REFERENCES

- [1] A. Shojaei and I. Flood, "Extending the Portfolio and Strategic Planning Horizon by Stochastic Forecasting of Unknown Future Projects," in *The Seventh International Conference on Advanced Communications and Computation, INFOCOMP 2017*, pp. 64–69.
- [2] B. S. Blichfeldt and P. Eskerod, "Project portfolio management - There's more to it than what management enacts," *Int. J. Proj. Manag.*, vol. 26, no. 4, pp. 357–365, May 2008.
- [3] J. A. Araúzo, J. Pajares, and A. Lopez-Paredes, "Simulating the dynamic scheduling of project portfolios," *Simul. Model. Pract. Theory*, vol. 18, no. 10, pp. 1428–1441, Nov. 2010.
- [4] R. G. Cooper, S. J. Edgett, and E. J. Kleinschmidt, "Portfolio management in new product development: Lessons from the leaders—I," *Res. Manag.*, vol. 40, no. 5, pp. 16–28, 1997.
- [5] H. Markowitz, "PORTFOLIO SELECTION*," *J. Finance*, vol. 7, no. 1, pp. 77–91, Mar. 1952.
- [6] W. F. McFarlan, "Portfolio approach to information systems," *Harv. Bus. Rev.*, vol. 59, no. 5, pp. 142–150, 1981.

- [7] T. R. Browning and A. A. Yassine, "Resource-constrained multi-project scheduling: Priority rule performance revisited," *Int. J. Prod. Econ.*, vol. 126, no. 2, pp. 212–228, Aug. 2010.
- [8] A. F. Carazo, T. Gómez, J. Molina, A. G. Hernández-Díaz, F. M. Guerrero, and R. Caballero, "Solving a comprehensive model for multiobjective project portfolio selection," *Comput. Oper. Res.*, vol. 37, no. 4, pp. 630–639, Apr. 2010.
- [9] M. Engwall, "No project is an island: Linking projects to history and context," *Res. Policy*, vol. 32, no. 5, pp. 789–808, May 2003.
- [10] B. Wang and Y. Song, "Reinvestment Strategy-Based Project Portfolio Selection and Scheduling with Time-Dependent Budget Limit Considering Time Value of Capital BT - Proceedings of the 2015 International Conference on Electrical and Information Technologies for Rail Transportat," 2016, pp. 373–381.
- [11] B. Canbaz and F. Marle, "Construction of project portfolio considering efficiency, strategic effectiveness, balance and project interdependencies," *Int. J. Proj. Organ. Manag.*, vol. 8, no. 2, p. 103, 2016.
- [12] M. Lehnert, A. Linhart, and M. Röglinger, "Value-based process project portfolio management: integrated planning of BPM capability development and process improvement," *Bus. Res.*, vol. 9, no. 2, pp. 377–419, Aug. 2016.
- [13] V. Mohagheghi, S. M. Mousavi, B. Vahdani, and M. R. Shahriari, "R&D project evaluation and project portfolio selection by a new interval type-2 fuzzy optimization approach," *Neural Comput. Appl.*, vol. 28, no. 12, pp. 3869–3888, Dec. 2017.
- [14] F. Faezy Razi and S. Hooman Shariat, "A hybrid grey based artificial neural network and C&R tree for project portfolio selection," *Benchmarking An Int. J.*, vol. 24, no. 3, pp. 651–665, 2017.
- [15] M. Shariatmadari, N. Nahavandi, S. H. Zegordi, and M. H. Sobhiyah, "Integrated resource management for simultaneous project selection and scheduling," *Comput. Ind. Eng.*, vol. 109, pp. 39–47, 2017.
- [16] J. M. Hummel, M. D. Oliveira, C. A. B. e Costa, and M. J. IJzerman, "Supporting the Project Portfolio Selection Decision of Research and Development Investments by Means of Multi-Criteria Resource Allocation Modelling," in *Multi-Criteria Decision Analysis to Support Healthcare Decisions*, Springer, 2017, pp. 89–103.
- [17] S.-S. Liu and C.-J. Wang, "Optimizing project selection and scheduling problems with time-dependent resource constraints," *Autom. Constr.*, vol. 20, no. 8, pp. 1110–1119, Dec. 2011.
- [18] W. R. Scott, *Organizations: Rational, Natural, and Open Systems*, 5th Editio. Prentice Hall, 2002.
- [19] M. Martinsuo, "Project portfolio management in practice and in context," *Int. J. Proj. Manag.*, vol. 31, no. 6, pp. 794–803, Aug. 2013.
- [20] R. B. Duncan, "Characteristics of Organizational Environments and Perceived Environmental Uncertainty," *Adm. Sci. Q.*, pp. 313–327, 1972.
- [21] R. L. Daft, "Organziation Theory and Design," in *South-Western Cengage Learning*, 2009, pp. 138–157.
- [22] M. M. Farshchian and G. Heravi, "Probabilistic Assessment of Cost, Time, and Revenue in a Portfolio of Projects Using Stochastic Agent-Based Simulation," *J. Constr. Eng. Manag.*, vol. 144, no. 5, p. 04018028, May 2018.
- [23] Y. Petit and B. Hobbs, "Project portfolios in dynamic environments: Sources of uncertainty and sensing mechanisms," *Proj. Manag. J.*, vol. 41, no. 4, pp. 46–58, Sep. 2010.
- [24] S. M. Shahandashti and B. Ashuri, "Highway Construction Cost Forecasting Using Vector Error Correction Models," *J. Manag. Eng.*, vol. 32, no. 2, p. 04015040, Mar. 2016.
- [25] A. Shojaei and I. Flood, "Stochastic forecasting of project streams for construction project portfolio management," *Vis. Eng.*, vol. 5, no. 1, p. 11, 2017.
- [26] A. Shojaei and I. Flood, "Stochastic Forecasting of Unknown Future Project Streams for Strategic Portfolio Planning," in *Computing in Civil Engineering 2017*, 2017, pp. 280–288.
- [27] C. Cargnoni, P. Müller, and M. West, "Bayesian Forecasting of Multinomial Time Series through Conditionally Gaussian Dynamic Models," *J. Am. Stat. Assoc.*, vol. 92, no. 438, pp. 640–647, Jun. 1997.
- [28] C. Voyant, G. Notton, C. Darras, A. Fouilloy, and F. Motte, "Uncertainties in global radiation time series forecasting using machine learning: The multilayer perceptron case," *Energy*, vol. 125, pp. 248–257, Apr. 2017.
- [29] J. Li and W. Chen, "Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models," *Int. J. Forecast.*, vol. 30, no. 4, pp. 996–1015, Oct. 2014.
- [30] P. Exterkate, P. J. F. Groenen, C. Heij, and D. van Dijk, "Nonlinear forecasting with many predictors using kernel ridge regression," *Int. J. Forecast.*, vol. 32, no. 3, pp. 736–753, Jul. 2016.
- [31] X. Yu and S.-Y. Liang, "Forecasting of hydrologic time series with ridge regression in feature space," *J. Hydrol.*, vol. 332, no. 3–4, pp. 290–302, Jan. 2007.
- [32] B. Choubin, S. Khalighi-Sigaroodi, A. Malekian, and Ö. Kişi, "Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals," *Hydrol. Sci. J.*, vol. 61, no. 6, pp. 1001–1009, Apr. 2016.
- [33] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. Neural Networks*, vol. 14, no. 6, pp. 1506–1518, Nov. 2003.
- [34] S. Thomas Ng, S. O. Cheung, R. Martin Skitmore, K. C. Lam, and L. Y. Wong, "Prediction of tender price index directional changes," *Constr. Manag. Econ.*, vol. 18, no. 7, pp. 843–852, Oct. 2000.
- [35] J. M. W. Wong and S. T. Ng, "Forecasting construction tender price index in Hong Kong using vector error correction model," *Constr. Manag. Econ.*, vol. 28, no. 12, pp. 1255–1268, 2010.

Progressive Advancement of Knowledge Resources and Mining: Integrating Content Factor and Comparative Analysis Methods for Dynamical Classification and Concordances

Claus-Peter Rückemann

Westfälische Wilhelms-Universität Münster (WWU),
Leibniz Universität Hannover,
North-German Supercomputing Alliance (HLRN), Germany
Email: ruckema@uni-muenster.de

Abstract—The research presented in this paper concentrates on the results from creating new advanced methodologies used for enhancing knowledge resources and knowledge mining. Creating and developing advanced knowledge resources and features for mining over long-time periods are challenging tasks, which require the continuous development of advanced complex means. Enhancements have to include semi-manual and automatable implementations from advanced methodologies in order to access the knowledge- and context-related characteristics and context values. This research builds on the practice of creation and development of multi-disciplinary knowledge resources for decades and creating and applying mining and discovery methods. The results on creating means for a comparative analysis of data entities from knowledge resources are used for data analysis and enhancement of complex and increasing multi-disciplinary knowledge resources. Comparing data entities is a most ambitious task for increasingly complex data objects, integrated resources, and relations – from the Knowledge Resources, as well as from the computational perspective. The implementation utilises complementary components, which enable to structure and describe complex knowledge and support an advanced analysis. Based on the methodological fundament, the paper presents practical results, delivers and discusses instructive examples from an implementation and case study. For practical reasons with the comparative analysis, the knowledge resources explicitly utilise references to publicly available data resources. The goal of this research is the enhancement of knowledge resources and mining by modular implementations of advanced methodologies, especially method integration for comparative analysis and knowledge mining with information systems and long-term multi-disciplinary knowledge resources.

Keywords—*Enhancement of Knowledge Resources; Comparative Analysis; Content Factor; Universal Decimal Classification; Advanced Data-centric Computing.*

I. INTRODUCTION

It is a truth universally acknowledged, that knowledge and knowledge-related data are core values of human activities. Knowledge and long-term knowledge resources are treasure troves of documentation and sources of major insights. The enhancement of knowledge resources and documentation also contributes to new discovery and insights. Therefore, new advanced methods need to be created contributing to the development and enhancement of knowledge resources. The fundaments of the method providing a Comparative Analysis of data

entities were presented at the INFOCOMP 2017 conference in Venice, Italy [1]. This paper presents results of extended research, which showed to be relevant for the enhancement of knowledge resources and mining and discovery processes. The enhancements include further aspects for integration of resources, the integration of conceptual knowledge, the architecture of the implementation, including the integration of different categories of resources, and computational aspects.

Advanced methods of knowledge mining with information systems and knowledge resources are becoming increasingly important. With that, improving knowledge mining and at the same time integrating larger amounts of data increases the challenges. The core of challenges is the data analysis. Within data analysis, comparing “data” is a central task. Comparing data entities is an even more ambitious task when data objects and relations are becoming more and more complex.

The term data entity in context with knowledge resources refers to any data representing objects of any kind like digital or realia objects, including references, e.g., to objects or conceptual knowledge. Within this research, special application components were created and implemented in order to provide modular means to be integrated for a comparative analysis, e.g., knowledge resources referring to structured and unstructured data, conceptual data, especially knowledge classification, and methods specialised on the before mentioned means, e.g., the Content Factor method (CONFACT) [2]. It is not the task of this practical implementation to re-iterate the basics of the instruments used. The basics of the method are explained in theory and practice in the original publication as well as theory, definitions, and context are explained and defined in referred publications.

The multi-disciplinary knowledge resources and the application of the Content Factor method have enabled new flexible workflows and the creation of new complementary means for both the enhancement of multi-disciplinary knowledge resources and for data-centric knowledge mining and discovery processes [3]. Some of the most widely required means with data entities of knowledge resources are components for a comparative analysis. Comparative Analysis (CA) is defined as an item-by-item comparison of two or more comparable entities. The goal of this research is the enhancement of knowledge resources and mining by appropriate integration of methods.

Resulting from this research, the enhancements are based on the new and original integration of Comparative Analysis and the Content Factor methods, an integration which is implemented and approved with the research projects in various practical context.

This paper is organised as follows. Section II summarises the state-of-the-art, motivation, and frame of reference to the ground of comparison. Section III introduces to knowledge and integration of different resources, as implemented for this case study. Section IV presents the integration of conceptual knowledge and Sections V and VI show the knowledge-centric integrative architecture for the computation and analysis of results based on the selected resources and the implementation of the workflow. Section VII discusses the main results and evaluates them in context of the application. Section VIII summarises the results and lessons learned, conclusions, and future work.

II. STATE-OF-THE-ART, MOTIVATION, AND FRAME

The elementary way of knowledge mining, practised by the vast majority of approaches and services ignores content quality, document types, and cognitive knowledge. That means, content is handled independently from the creation process and expertise, content from databases, Web pages, and scanned books are not differentiated, and classification of content is disregarded.

CA modules can be used for arbitrary purposes with knowledge mining workflows, e.g., for selecting complementary resources as well as selecting objects supporting decision making processes. The method is used with knowledge mining workflows, integrating dedicated knowledge resources and publicly available content, e.g., text documents and books, because of their complementary nature regarding content, structure, and quality.

The complexity requires to start with a comprehensive high level view and context for the target of this research. The following sections describe the motivation and the base of the conducted CA.

A. Frame of reference

The significance of integrating different data entities results from the context, in which they are placed. This research presents a method of comparing different data entities as referred from objects in advanced knowledge resources.

Objects with higher quality are mostly more complex. Advanced knowledge mining and decision making requires more than one method or algorithm for analysis of available objects and their references, data entities, and attributes. A major challenge is the difference of entities, e.g., regarding entity type, original purpose of the entity, and source but also content and structure.

Different types of entities cannot be ignored from advanced workflows because they contain unique knowledge and information. In most cases, the knowledge and information can even only be provided by different entities and referred sources. Methods should be provided, which are beneficial to be integrated in advanced workflows, especially for analysis,

quantisation, and qualification of different entities. The deployed means should allow long-term data-centric applications and intrinsically foster the seamless integration with existing workflows. In addition, the methodologies, methods, and architecture of integration should allow the implementation of modular and least invasive components.

B. Grounds for comparison

Besides the complexity, a combination of data entities from different sources and different types was chosen for the following reasons. The rationale behind the choice for knowledge resources and entities from referred objects results from complementary content and context. There is an arbitrary high quality of multi-disciplinary content in the knowledge resources, which are in continuous development [4]. In addition, the knowledge resources can provide an extremely high knowledge and information density. The Gutenberg resources [5] can provide a large number of fully publicly available standard text documents and elaborations for a wide multi-disciplinary context. Both types of resources contain essential amounts of textual content and are continuously extended and improved. The relationship between different entities is the addressed knowledge content with its unique nature. The thesis is, that different entities should neither be left out from advanced workflows nor should their content, the unique knowledge and information, be ignored.

C. Organisational scheme

The targeted lens comparison discusses the most important aspects text-by-text, focussing on advanced knowledge resources and referred resources.

For complex lens comparisons of that kind we require to have an extended focus on integration of resources, including aspects of automatically created and integrated objects. The integration further has to resolve and manage conceptual knowledge, especially for important sources of conceptual knowledge. Creating concordances is a common means used with classifications. Concordances are used for providing references between classifications, they are not used for comparing or analysing content or data. With the overall goal of enhancing knowledge resources and mining, supporting and generating concordances are major means for achieving the integration of conceptual knowledge. Special care is taken for aspects of computation and analysis, due to the large complexity and data and the requirements for compute intensive advanced methodologies and algorithms.

III. KNOWLEDGE AND THE INTEGRATION OF RESOURCES

The following sections describe how an integration was achieved and which results were gained with the analysis.

A. Knowledge

One of the most important sources of understanding knowledge is Aristoteles' treatise of the Nicomachean Ethics [6], which is a classic basis of its essence [7], significance, and terms [8].

If one re-visits a place it will, to some extent, be a different place [9]. There are long-term and short term changes. Everything in the world is connected. This is also true for knowledge resources, embedded in existing context. Knowledge resources contain multi-disciplinary knowledge objects, which can be used in arbitrary ways for providing knowledge, e.g., factual, conceptual, procedural, and metacognitive knowledge [10].

The objects [11] can contain any content and context as well as references, e.g., translations, transliterations, synonyms, associations [12], references [13], conceptual knowledge (e.g., UDC), concordances [14], links, references (see, “s.”), optional references (see also, “s. also”), comparable references (compare, “comp.”), keywords, and Content Factors (of elements) [15]. The objects can be based on records, e.g., characters, words, lines, and complex records. In practical application scenarios [16], any content and context can be used for analysis and evaluation of an object.

B. Data entities

Data entities can be created from many resources. With this research, knowledge objects and data entities were automatically created from ‘Gutenberg documents’. At the time of the case study (January 2017) Project Gutenberg [5] offered 53,855 free ebooks for download. At the time of preparation of this article (February 2018) Project Gutenberg offered 56,432 free ebooks for download. The document files include the text in a version of the respective edition, which can be a revised edition or translation. The text editions are linked as different document files, e.g., plain text files, which can be converted into data entities and integrated with different data entities. Regarding conceptual knowledge, the Gutenberg documents use a flat implementation of the Library of Congress (LoC) classification outline [17]. The ebook links contain some relevant information, too, e.g., the bibliographic record, EBook-No. 25062, a link the LoC Class entries, and the release date of the edition. The original publication date of the source text is contained in the document files. Publications like books and articles are not “unstructured”. They differ in structure from knowledge resources’ content but they can be seen as structured entities, too. Data entities from knowledge resources’ collections and containers [2] are used with many knowledge mining applications [4].

With the created modules the data entities from the Gutenberg resources can be handled in the same way as knowledge resources’ objects, e.g., of different origin. The following case study starts with a knowledge mining request for “Vesuvius” in the context of “volcanology”. The primary Gutenberg result matrix contains a number of documents [18]-[24] provided with the precomputation. The essential steps and data of the examples should be comprehensible as these chosen resources are publicly available.

C. Object and data entity integration: Four cases

The following passages introduce four different types of objects, which were considered in order to integrate the available knowledge. The four types of objects are originating from three

major groups of resources, namely publicly provided book and document object, collection objects, and container objects. For the first group, it is shown that the respective objects can have significant different characteristics, which is shown comparing two samples.

Objects and data entities can be integrated with knowledge resources in arbitrary ways, e.g., as a referred object or by creating an instance of an object. Here, for the goal of this research, the programming and programming languages of the modules are not relevant for the demonstration. Significant is the integration, which allows an analysis and evaluation, e.g., with knowledge resources’ objects. The following excerpt (Figure 1) shows a knowledge resources’ object automatically created from an entity of Gutenberg document 33483 [19] with LoC classification [25].

```

1 33483-0.txt [Document]:
2 ...
3 THE
4 ERUPTION OF VESUVIUS
5 IN 1872, ...
6 BY
7 PROFESSOR LUIGI PALMIERI,
8 _Of the University of Naples; Director of the Vesuvian Observatory...
9 ...
10 WITH NOTES, AND AN
11 _INTRODUCTORY SKETCH OF THE PRESENT STATE OF KNOWLEDGE_
12 OF
13 TERRESTRIAL VULCANICITY,
14 _The Cosmical Nature and Relations of
15 Volcanoes and Earthquakes._
16 ...
17 BY
18 ROBERT MALLETT,
19 _Mem. Inst. C.E., F.R.S., F.G.S., M.R.I.A., &c., &c._
20 ...
21 WITH ILLUSTRATIONS. ...
22 LONDON: ...
23 _ASHER & CO._,
24 13, BEDFORD STREET, COVENT GARDEN, W.C. ...
25 1873. ...
26 W. S. Johnson, Nassau Steam Press, 60, St. Martin's Lane,
27 Charing Cross, W.C.
28 ...

```

Figure 1. Automatically created Gutenberg knowledge resources object for document 33483 (geosciences collection, LX, excerpt).

The following excerpt (Figure 2) shows a knowledge resources’ object automatically created from an entity of Gutenberg document 25062 [21] with LoC classification [25].

```

1 pg25062.txt [Document]:
2 ...
3 A STUDY OF
4 RECENT EARTHQUAKES.
5 ...
6 BY
7 CHARLES DAVISON, Sc.D., F.G.S.
8 ...
9 AUTHOR OF
10 "THE_HEREFORD_EARTHQUAKE_OF_DECEMBER_17TH,_1896."
11 ...
12 WITH 80 ILLUSTRATIONS
13 ...
14 London and Newcastle-on-Tyne:
15 THE WALTER SCOTT PUBLISHING CO., LTD.
16 1905
17 ...
18 PREFACE.
19 ...
20 The present volume differs from a text-book of seismology in giving
21 brief, though detailed, accounts of individual earthquakes rather
22 than
23 a discussion of the phenomena and distribution of earthquakes in
  general. ...

```

Figure 2. Automatically created Gutenberg knowledge resources object for document 25062 (geosciences collection, LX, excerpt).

Both objects share the same LoC classification. Without advanced means and further analysis both might be considered providing knowledge for the same topics and purposes.

On the other hand, as an example from a different category of resources, an object excerpt of an object instance “Vesuvius” from a knowledge resources’ collection, resulting from a knowledge mining process, is shown in Figure 3.

```

1 Vesuvius [Volcanology, Geology, Archaeology]:
2 (lat.) Mons Vesuvius.
3 (ital.) Vesuvio.
4 Volcano, Gulf of Naples, Italy.
5 Complex volcano (compound volcano).
6 Stratovolcano, large cone (Gran Cono).
7 Volcano Type: Somma volcano,
8 VNUM: 0101-02=,
9 Summit Elevation: 1281\UD(m). ...
10 ...
11 Syn.: Vesaevus, Vesevus, Vesbius, Vesvius
12 s. volcano, super volcano, compound volcano
13 s. also Pompeji, Herculaneum, seismology
14 ...
15 compare La Soufrière, Mt. Scenery, Soufriere
16 ...
17 %%IML: UDC:[911.2+55]:[57+930.85]:[902]*63*(4+37+23+24)=12=14
18 %%IML: GoogleMapsLocation: http://maps.google.de/maps?hl=de&gl=de&vpsrc
   =0&ie=UTF8&ll=40.821961,14.428868&spn=0.018804,0.028238&t=h&z=15
19 ...
20 ...
21 ...
22 ...
23 ...
24 ...

```

Figure 3. Knowledge resources collection object “Vesuvius” (LX resources, geoscientific collection, excerpt).

The objects can contain any knowledge, e.g., factual and conceptual knowledge. Here, the object carries names and synonyms in different languages, dynamically usable geocoordinates, Universal Decimal Classification (UDC) and so on, including geoclassification (UDC:(37), Italia. Ancient Rome and Italy). In addition to collection objects, another important source of knowledge are container type resources, which contain objects, mostly highly comparable from perspectives of content and structure. Figure 4 shows a tiny excerpt of a processed volcanological features container.

```

1 UCC:UDC2012:551.21
2 UCC:UDC2012:551
3 UCC:UDC2012:551.2,551.23,551.24,551.26
4 UCC:UDC2012:902/908
5 UCC:MSC2010:86,86A17,86A60
6 UCC:LCC:QE521-545
7 UCC:LCC:QE1-996.5
8 UCC:LCC:QC801-809
9 UCC:LCC:CCL-960,CB3-482
10 UCC:PACS2010:91.40.-k
11 UCC:PACS2010:91.65.-n,91.
12 UCC:PACS2010:91.40.Ge,91.40.St,91.40.Rs,+91.45.C-,+91.45.D-,90
13 ...
14 CONTAINER_OBJECT_EN_ITEM: Vesuvius
15 CONTAINER_OBJECT_DE_ITEM: Vesuv
16 CONTAINER_OBJECT_EN_PRINT: Vesuvius
17 CONTAINER_OBJECT_DE_PRINT: Vesuv
18 CONTAINER_OBJECT_EN_COUNTRY: Italy
19 CONTAINER_OBJECT_DE_COUNTRY: Italien
20 CONTAINER_OBJECT_EN_CONTINENT: Europe
21 CONTAINER_OBJECT_DE_CONTINENT: Europa
22 CONTAINER_OBJECT_XX_LATITUDE: 40.821N
23 CONTAINER_OBJECT_XX_LONGITUDE: 14.426E
24 CONTAINER_OBJECT_XX_HEIGHT_M: 1281
25 CONTAINER_OBJECT_EN_TYPE: Complexvolcano
26 CONTAINER_OBJECT_DE_TYPE: Komplex-Vulkan
27 CONTAINER_OBJECT_XX_VNUM: 0101-02=

```

Figure 4. Knowledge resources container object, processed instance of a simple container entry “Vesuvius” (LX resources, excerpt).

The container objects comprise of various knowledge, especially factual and conceptual references, including concordances. Objects can also contain multi-lingual entries. The organisation of the objects in a specific container is similar to identical. The resources’ access and processing can be done in any programming language, assuming that the interfaces are

implemented. For example, combining scripting, filtering, and parallel programming can provide flexible approaches.

The data used here is based on the content and context from the knowledge resources, provided by the LX Foundation Scientific Resources (LX not an acronym) [4]. The integration and implementation of conceptual knowledge and concordances as shown part of the objects will be discussed in the next section.

IV. INTEGRATION OF CONCEPTUAL KNOWLEDGE

Conceptual knowledge (e.g., classification) is a very important means for describing objects. For the Gutenberg resources a simple LoC classification for every document is provided, which allows to get the major topic classification.

The knowledge resources can use any classification for describing objects, elements, and views. This also allows to use multiple classifications and even specialised classifications for the description. The resources can also provide concordances for mapping the classifications.

All these means can be used for the documentation and analysis of objects, e.g., comparisons. The conceptual knowledge provides the facilities to handle objects and entities even from different sources and context.

A. Implemented references to conceptual knowledge

The objects resulting from the mining request for “Vesuvius” in the context of “volcanology” are also referred with conceptual knowledge. Table I shows some examples of Gutenberg objects and their classification referenced with the knowledge resources.

TABLE I. GUTENBERG CONCEPTUAL KNOWLEDGE: LOC.

Ref.	LoC Code and Description	LoC Ref.
[18]	QE: Science: Geology	[25]
[19]	DH: History: General and Eastern Hemisphere: Netherlands, Belgium, Luxemburg	[26]
[20]	QB: Science: Astronomy	[27]
[21]	QE: Science: Geology	[25]
[28]	PS: Language and Literatures: American and Canadian literature	[29]
[22]	QE: Science: Geology	[25]
[23]	QE: Science: Geology	[25]
[24]	QE: Science: Geology	[25]

A means of integration with other classifications is the use of concordances. Table II shows the according excerpt of concordances with LoC and UDC.

TABLE II. LOC TO UDC CONCORDANCES.

LoC Code	UDC Code	UDC Reference
QE: ...	UDC:55	[30]
DH: ...	UDC:93/94	[31][32][33]
QB: ...	UDC:52	[34]
QE: ...	UDC:55	[30]
PS: ...	UDC:821.111	[35]
QE: ...	UDC:55	[30]
QE: ...	UDC:55	[30]
QE: ...	UDC:55	[30]

Classifications of UDC editions consistently refer to verbal descriptions. For this part of the research all small unsorted excerpts of the knowledge resources objects only refer to main UDC-based classes, which for this part of the publication are taken from the Multilingual Universal Decimal Classification Summary (UDCC Publication No. 088) [36] released by the UDC Consortium under the Creative Commons Attribution Share Alike 3.0 license [37] (first release 2009, subsequent update 2012). Table III lists the UDC classifications and their verbal description in English.

TABLE III. UDC RESOLVED, VERBAL DESCRIPTION ENGLISH.

UDC Code	Description	Ref.
UDC:55	Earth Sciences. Geological sciences	[30]
UDC:93/94	History	[31][32][33]
UDC:52	Astronomy. Astrophysics. Space research. Geodesy	[34]
UDC:55	Earth Sciences. Geological sciences	[30]
UDC:821.111	English literature	[35]
UDC:55	Earth Sciences. Geological sciences	[30]
UDC:55	Earth Sciences. Geological sciences	[30]
UDC:55	Earth Sciences. Geological sciences	[30]

The table lists those entries resulting from the concordances and from associated objects. The UDC is available in about 50 languages. The English verbal descriptions were used for this case study. The verbal description can be included in the context creation and can, for example, provide scalable fuzziness for creating multi-disciplinary context. The LX knowledge resources' structure and the classification references [38] based on UDC [39], [40], [36] are essential means for the processing workflows and evaluation of the knowledge objects and containers. Both provide strong multi-disciplinary and multi-lingual support.

B. Concordances

Different resources as well as different disciplines may use different classifications. Gutenberg is currently using the LoC. The knowledge resources are using a Universal Classified Classification (UCC) both with classification and concordances for the objects collected over time. The listing in Figure 5 shows a simple example for concordances.

```

1  ...
2  UCC:UDC:55
3  UCC:LCC:QE
4
5  ...
6  UCC:UDC:93/94
7  UCC:LCC:DH
8
9  ...
10 UCC:UDC:52
11 UCC:LCC:QB
12
13 ...
14 UCC:UDC:821.111
15 UCC:LCC:PS
16
17 ...

```

Figure 5. Classification and concordances excerpt of a simple object instance (knowledge resources collection).

The references to individual, specialised and universal classifications consistently describe the conceptual knowledge as correct as possible within a classification.

The concordances integrate the context of more than one classification. This enhances facilities for in-depth description and integration, the specialisation on conceptual knowledge as well as the broadness of universal knowledge and context. The listing in Figure 6 shows a simple object instance classification and concordances excerpt from a volcanological object in a collection.

```

1  ...
2  UCC:UDC2012:551.21
3  UCC:UDC2012:551
4  UCC:UDC2012:902/908
5  UCC:MSC2010:86,86A17,86A60
6  UCC:LCC:QE521-545
7  UCC:LCC:QE1-996.5
8  UCC:LCC:QC801-809
9  UCC:LCC:CC1-960,CB3-482
10 UCC:PACS2010:91.40.-k
11 UCC:PACS2010:91.65.-n,91.

```

Figure 6. Classification and concordances excerpt of a simple object instance (knowledge resources collection).

The excerpt shows classification concordances in several different classifications as used in different disciplines. Even a lot of internal details of such concordances are self-explanatory with a basic knowledge and practice of the used classifications. Concordances also interlink different classifications and disciplines. To a certain extent most classifications also express the context of certain disciplines. Possibly multiple views from different disciplines or author groups on a certain object are not shown in this reduced view but they can also hold the full spectrum of classifications and concordances and also express views and development of views and object instances over time.

V. KNOWLEDGE-CENTRIC INTEGRATIVE ARCHITECTURE

The analysis of integrated resources requires advanced methods and algorithms. A method used for description and analysis of objects is the Content Factor.

A. Content Factor computation for data entities

Objects of any kind can be integrated with knowledge resources. Objects can contain instances of data entities and refer to associated knowledge. For an analysis, a number of common information regarding the objects and data entities is required.

Examples were shown in the previous section on object and data entity integration, where objects were transformed into collection objects, including references to further knowledge like factual knowledge and classification.

The following excerpt (Figure 7) illustrates the creation of Content Factor definition sets [2] for the use with data entities. All the basics details and the algorithm of the method are described in the original publication. Definition sets are used for both Gutenberg resources and knowledge resources.

```

1 % (c) LX-Project, 2016, 2017
2 {Vol}:= [Vv] [Oo] [Ll] [Cc] [Aa] [Nn]
3 {Vul}:= [Vv] [Uu] [Ll] [Cc] [Aa] [Nn]
4 {Veu}:= [Vv] [Ee] [Ss] [Uu] [Vv]
5 {Vee}:= [Vv] [Ee] [Ss] [Ee] [Vv] [Oo]
6 {Kom}:= [Kk] [Oo] [Mm] [Ee] [Tt]
7 {Com}:= [Cc] [Oo] [Mm] [Ee] [Tt]
8 {Met}:= [Mm] [Ee] [Tt] [Ee] [Oo] [Rr]
9 {Erd}:= [Ee] [Rr] [Dd] [Bb] [Ee] [Bb] [Ee] [Nn]
10 {Ear}:= [Ee] [Aa] [Rr] [Tt] [Hh] [Qq] [Uu] [Aa] [Kk] [Ee]
11 {Puz}:= [Pp] [Uu] [Zz] [Zz] [Oo] [Ll]
12 {Poz}:= [Pp] [Oo] [Zz] [Zz] [Oo] [Ll]
    
```

Figure 7. CONTFACT definition set for Gutenberg Project resources and knowledge resources, (LX, excerpt).

Figure 8 shows the Normed Basic Content Factor (NBCF, $\bar{\kappa}_B$) [2] computed for a knowledge resources object reference to the Gutenberg Project document 33483.

```

1 CONTFACT:BEGIN
2 CONTFACT:20161227-234624:AU: {Veu} {Vul} {Vol} {Ear} {Veu} {Met} {Veu} {Vol} {Met}
3 ... {Veu} {Veu} {Ear} {Veu} ... {Veu} {Veu} {Veu} {Veu} ... {Veu} {Ear} {Vol} {Ear} /39843
4 CONTFACT:20161227-234624:AS: {Ear} ... {Veu} {Veu} {Vol} {Vol} ... {Vul} {Vul} /39843
5 CONTFACT:20161227-234624:M: {Vul}:= [Vv] [Uu] [Ll] [Cc] [Aa] [Nn]
6 CONTFACT:20161227-234624:M: {Veu}:= [Vv] [Ee] [Ss] [Uu] [Vv]
7 CONTFACT:20161227-234624:M: {Vee}:= [Vv] [Ee] [Ss] [Ee] [Vv] [Oo]
8 CONTFACT:20161227-234624:M: {Kom}:= [Kk] [Oo] [Mm] [Ee] [Tt]
9 CONTFACT:20161227-234624:M: {Com}:= [Cc] [Oo] [Mm] [Ee] [Tt]
10 CONTFACT:20161227-234624:M: {Met}:= [Mm] [Ee] [Tt] [Ee] [Oo] [Rr]
11 CONTFACT:20161227-234624:M: {Erd}:= [Ee] [Rr] [Dd] [Bb] [Ee] [Bb] [Ee] [Nn]
12 CONTFACT:20161227-234624:M: {Ear}:= [Ee] [Aa] [Rr] [Tt] [Hh] [Qq] [Uu] [Aa] [Kk] [Ee]
13 CONTFACT:20161227-234624:M: {Puz}:= [Pp] [Uu] [Zz] [Zz] [Oo] [Ll]
14 CONTFACT:20161227-234624:M: {Poz}:= [Pp] [Oo] [Zz] [Zz] [Oo] [Ll]
15 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSDEF=11
16 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSALL=39843
17 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSMAT=356
18 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSCFO=.00900324
19 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSKWO=1
20 CONTFACT:20161227-234624:M:STAT:OBJECTELEMENTSLAN=0
21 CONTFACT:20161227-234624:M:INFO:OBJECTELEMENTSOBJ=33483-0.txt
22 CONTFACT:20161227-234624:M:INFO:OBJECTELEMENTSDCM=(c) LX-Project, 2016, 2017
23 CONTFACT:20161227-234624:M:INFO:OBJECTELEMENTSMTX=LX Foundation Scientific
24 Resources; Object Collection
25 CONTFACT:20161227-234624:M:INFO:OBJECTELEMENTSAUT=Claus-Peter R\*uckemann
26 CONTFACT:END
    
```

Figure 8. NBCF $\bar{\kappa}_B$ computed for knowledge resources object reference to Gutenberg Project document 33483 (LX Resources, excerpt).

Figure 9 shows the Normed Basic Content Factor (NBCF, $\bar{\kappa}_B$) computed for a knowledge resources object reference to the Gutenberg Project document 25062.

```

1 CONTFACT:BEGIN
2 CONTFACT:20161227-234626:AU: {Ear} ... {Vol} {Ear} ... {Veu} {Vol} {Ear} ... {Veu} {Ear} {
3 Ear} {Ear} {Veu} ... {Veu} ... {Veu} {Ear} ... {Ear} /88463
4 CONTFACT:20161227-234626:AS: {Ear} ... {Veu} ... {Vol} {Vul} {Vul} /88463
5 CONTFACT:20161227-234626:M: {Vul}:= [Vv] [Oo] [Ll] [Cc] [Aa] [Nn]
6 CONTFACT:20161227-234626:M: {Veu}:= [Vv] [Ee] [Ss] [Uu] [Vv]
7 CONTFACT:20161227-234626:M: {Vee}:= [Vv] [Ee] [Ss] [Ee] [Vv] [Oo]
8 CONTFACT:20161227-234626:M: {Kom}:= [Kk] [Oo] [Mm] [Ee] [Tt]
9 CONTFACT:20161227-234626:M: {Com}:= [Cc] [Oo] [Mm] [Ee] [Tt]
10 CONTFACT:20161227-234626:M: {Met}:= [Mm] [Ee] [Tt] [Ee] [Oo] [Rr]
11 CONTFACT:20161227-234626:M: {Erd}:= [Ee] [Rr] [Dd] [Bb] [Ee] [Bb] [Ee] [Nn]
12 CONTFACT:20161227-234626:M: {Ear}:= [Ee] [Aa] [Rr] [Tt] [Hh] [Qq] [Uu] [Aa] [Kk] [Ee]
13 CONTFACT:20161227-234626:M: {Puz}:= [Pp] [Uu] [Zz] [Zz] [Oo] [Ll]
14 CONTFACT:20161227-234626:M: {Poz}:= [Pp] [Oo] [Zz] [Zz] [Oo] [Ll]
15 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSDEF=11
16 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSALL=88463
17 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSMAT=986
18 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSCFO=.01122068
19 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSKWO=1
20 CONTFACT:20161227-234626:M:STAT:OBJECTELEMENTSLAN=0
21 CONTFACT:20161227-234626:M:INFO:OBJECTELEMENTSOBJ=pq25062.txt
22 CONTFACT:20161227-234626:M:INFO:OBJECTELEMENTSDCM=(c) LX-Project, 2016, 2017
23 CONTFACT:20161227-234626:M:INFO:OBJECTELEMENTSMTX=LX Foundation Scientific
24 Resources; Object Collection
25 CONTFACT:20161227-234626:M:INFO:OBJECTELEMENTSAUT=Claus-Peter R\*uckemann
26 CONTFACT:END
    
```

Figure 9. NBCF $\bar{\kappa}_B$ computed for knowledge resources object reference to Gutenberg Project document 25062 (LX Resources, excerpt).

Figure 10 shows the NBCF computed for a knowledge resources object reference to the object “Vesuvius” (Figure 3).

```

1 CONTFACT:BEGIN
2 CONTFACT:20170205-161508:AU: {Veu} {Vol} {Veu} {Veu} {Vol} {Vol} {Vol} {Vol} {Vol} {
3 Vol} {Vol} {Vol} /71
4 CONTFACT:20170205-161508:AS: {Veu} {Veu} {Veu} {Vol} {Vol} {Vol} {Vol} {Vol} {Vol} {
5 Vol} {Vol} {Vol} /71
6 CONTFACT:20170205-161508:M: {Vul}:= [Vv] [Oo] [Ll] [Cc] [Aa] [Nn]
7 CONTFACT:20170205-161508:M: {Vul}:= [Vv] [Uu] [Ll] [Cc] [Aa] [Nn]
8 CONTFACT:20170205-161508:M: {Veu}:= [Vv] [Ee] [Ss] [Uu] [Vv]
9 CONTFACT:20170205-161508:M: {Vee}:= [Vv] [Ee] [Ss] [Ee] [Vv] [Oo]
10 CONTFACT:20170205-161508:M: {Kom}:= [Kk] [Oo] [Mm] [Ee] [Tt]
11 CONTFACT:20170205-161508:M: {Com}:= [Cc] [Oo] [Mm] [Ee] [Tt]
12 CONTFACT:20170205-161508:M: {Met}:= [Mm] [Ee] [Tt] [Ee] [Oo] [Rr]
13 CONTFACT:20170205-161508:M: {Erd}:= [Ee] [Rr] [Dd] [Bb] [Ee] [Bb] [Ee] [Nn]
14 CONTFACT:20170205-161508:M: {Ear}:= [Ee] [Aa] [Rr] [Tt] [Hh] [Qq] [Uu] [Aa] [Kk] [Ee]
15 CONTFACT:20170205-161508:M: {Puz}:= [Pp] [Uu] [Zz] [Zz] [Oo] [Ll]
16 CONTFACT:20170205-161508:M: {Poz}:= [Pp] [Oo] [Zz] [Zz] [Oo] [Ll]
17 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSDEF=11
18 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSALL=71
19 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSMAT=13
20 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSCFO=.21311475
21 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSKWO=2
22 CONTFACT:20170205-161508:M:STAT:OBJECTELEMENTSLAN=1
23 CONTFACT:20170205-161508:M:INFO:OBJECTELEMENTSOBJ=Vesuvius
24 CONTFACT:20170205-161508:M:INFO:OBJECTELEMENTSDCM=(c) LX-Project, 2016, 2017
25 CONTFACT:20170205-161508:M:INFO:OBJECTELEMENTSMTX=LX Foundation Scientific
26 Resources; Object Collection
27 CONTFACT:20170205-161508:M:INFO:OBJECTELEMENTSAUT=Claus-Peter R\*uckemann
28 CONTFACT:END
    
```

Figure 10. NBCF $\bar{\kappa}_B$ computed for knowledge resources object “Vesuvius” (LX Resources, excerpt).

All NBCF were computed with the same definition set (Figure 7). The data entities from the referenced Gutenberg resources and knowledge resources both contain multiple matches. The resulting Content Factor for the knowledge resources object is higher due to the higher concentration of relevant elements in the object. The Gutenberg object shows a higher absolute number of matches and multiple hits.

B. Architecture and integration

The Content Factor provides a measure, which can be used for describing characteristics of different objects.

The method can be used in general for any kind of object and is not restricted to limited kinds of documents. The Gutenberg document resources were chosen for the case study because of their range of content, type of documents, and availability of different text formats, which correspond with the intended integration with other resources.

The following diagram (Figure 11) depicts the architecture design of the conceptual solution as used for the implementation. The illustration shows the categories of resources and their integration regarding their contributions to a Comparative Analysis workflow, implemented for the use with the Gutenberg resources.

The workflow shows the abstract mathematical and computations details of the integration of resources, especially the interfaces, analysis, join, visualisation, statistics, and plotting as of the developed solution. With this research, resources can be divided in three major categories, symbolised as columns, depending on their nature and characteristics.

The central resources are knowledge resources, containing collections and containers as well as referenced and integrated resources. Content is arbitrary, e.g., any factual knowledge, conceptual knowledge, and procedural knowledge.

In this case, the enhancement of the means of mining and the knowledge resources are subject ordinary resources, especially the resources of the Gutenberg Project [5].

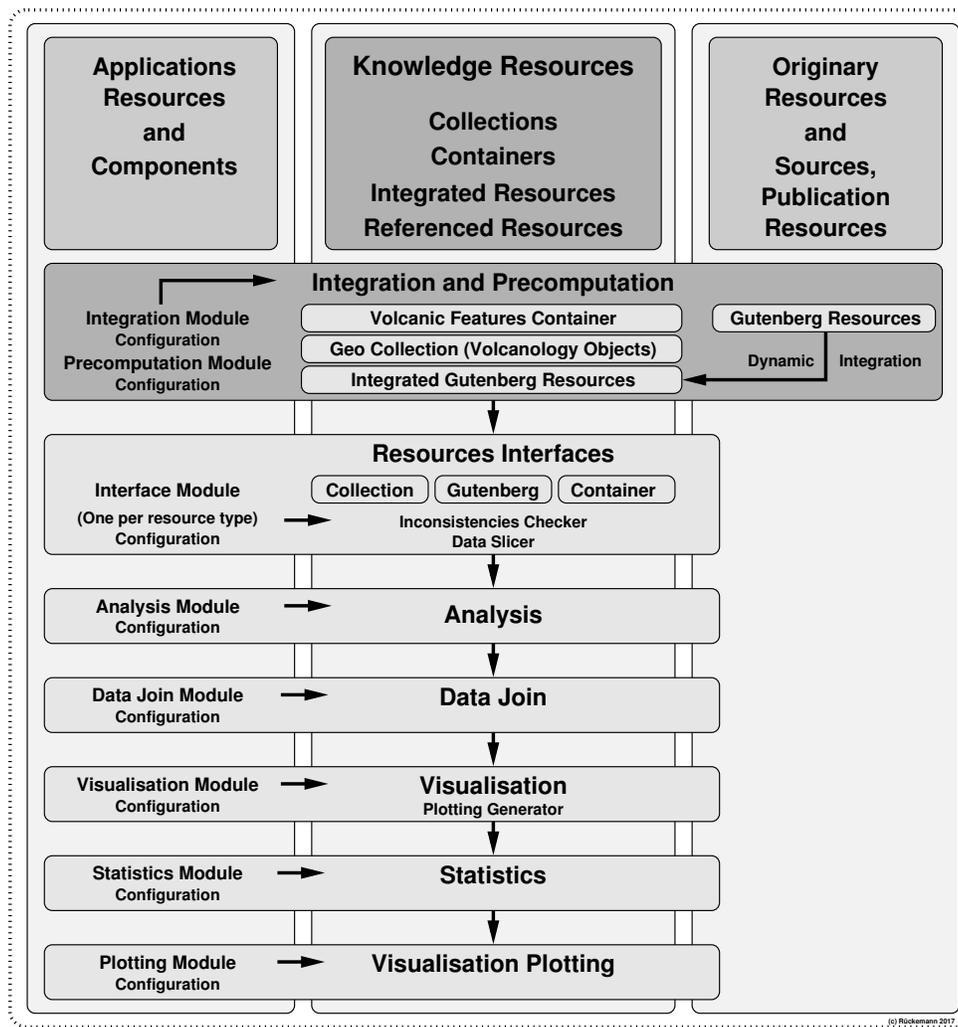


Figure 11. Architecture and integration: Categories of resources and their integration, showing the contributions to a Comparative Analysis workflow, implemented for the use with the Gutenberg resources being dynamically integrated, e.g., for analysis, visualisation, and statistics in an automated decision making workflow.

For the mining example, the integration and precomputation has to care for volcanic features container, geo collection, and the integrated objects from the Gutenberg resources.

C. Knowledge Resources and orinary resources

The analysis of different classifications, development of concepts for intermediate classifications, and experiences from case studies from the research conducted in the Knowledge in Motion (KiM) long-term project [41] have contributed to the application of UDC and different classifications and concordance schemes in the context of knowledge resources.

The following term definitions for object, container, and matrix can be helpful in this context.

- An object is an entity of knowledge data being part of knowledge resources. An object can contain any documentation, references, and other data. Objects can have an arbitrary number of sub-objects.

- A container is a collection of knowledge objects in a conjoint format.
- A matrix is a subset of the entirety, the “universe”, of knowledge. A workflow can consist of many sub-workflows each of which can be based on an arbitrary number of knowledge matrices. The output of any sub-workflow or workflow can be seen as an intermediate or final result matrix.

The flexible creation of objects carrying references, especially classification and concordances is the fundament for advanced knowledge processing and computing.

Collections can hold objects, which are more or less individual consist of any smaller entities and can have any references to other objects or resources. Containers can hold object, which have a comparable structure and comparable entities, e.g., objects belonging to a certain field of research or discipline. External resources can be referred from these objects and resources but it is also possible to integrate objects,

collections, and containers from external resources.

The category of ordinary resources can contain realia objects, original digital objects, as well as sources and publication resources, e.g., physical or digital books or proceedings. Most of the resources in this category restrict access to their content and provide limited interfaces and individual structures.

D. Computation and integration

The category of application resources and components can hold applicable implementations, e.g., software routines, interfaces, and services.

The Gutenberg resources were chosen for this implementation and the practical case study. The Gutenberg resources provide published materials for free Open Access, especially books, but neither provide interfaces nor further usable structures like containers in the aforementioned meaning.

The implementation therefore considered an integration of Gutenberg resource entities into the knowledge resources, on the level of collections and containers. The integration can be done dynamically and non-dynamically as well as the integration can be persistent with the knowledge resources or not. In any case, the reason is to create objects and entities, which are at a comparable level with other available objects, in content type and structure. The objects and entities are precomputed for that purpose. Therefore, with the application resources, the implementation can provide modules for the integration and precomputation. In consequence, the workflow (Figure 11) can rely on the integrated resources and the modules provided by the application resources.

The workflow requires interface modules for the participating resources, one individual module per resource type. This step includes additional preparation, e.g., inconsistencies checkers and data slicers. Data slicers are partitioning the available content in a way in which seems reasonable to have a common analysis with the integrated objects, e.g., based on full objects, entities, text blocks or lines.

The next steps do the analysis of objects, also considering their individual nature and original, in order to prepare for joining the data. A consequent step does an intermediate visualisation, generating a plotting routine for further consequent statistics, analysis and visualisation. The plotting generator is especially significant in this case study because the partial intermediate visualisation can deliver important information for further statistics and visualisation.

VI. IMPLEMENTATION OF THE WORKFLOW

A. Procedures and modules

Two main modules were required with the assistance pre-computation for identifying and selecting objects and data entities from the Gutenberg resources before entering the CA workflow. The preparative assistance data was computed with a module `gutenberganalysis` and the classification was extracted with a module `gutenbergloc`.

The first module extracts the desired content of an specific object and checks for association with the mining request, e.g., keywords and referenced knowledge. The second module extracts the classification of a specific object.

The CA workflow builds on these preparatory results. The implementation case study for the comparative analysis method required the creation of several major components and modules. Table IV shows a sequence of modules, which allows to create the base for a CA workflow as created with this case study.

TABLE IV. COMPARATIVE ANALYSIS WORKFLOW PROCEDURES AND IMPLEMENTED MODULES WITH GUTENBERG RESOURCES.

<i>Procedure</i>	<i>Module</i>
Gutenberg interface	<code>textca_gutenberginterface</code>
Configuration	
Inconsistencies checker	
Data slicer	
Analysis	<code>textca_analysis</code>
Configuration	
Data join	<code>textca_join</code>
Configuration	
Visualisation module	<code>textca_visualisation</code>
Configuration	
Plotting generator	
Conditional visualisation	
Statistics	<code>textca_statistics</code>
Configuration	
Visualisation plotting	<code>textca_plotting</code>
Configuration	

The workflow shows the principle procedures. Practically, the modules can be implemented with any environment and frameworks. In the case study Perl [42], Shell, and Gnuplot [43] were used. In general this means any module could be replaced by a different implementation separately.

Any module requires configuration options, which at least can be pre-configured options. In their application, the analysis up to visualisation modules for the Gutenberg resources are identical to the application for the knowledge resources. Therefore, the computations for all data entities were done with the `textca` group of modules. The module special for the Gutenberg resources is the Gutenberg interface, which requires to implement the proper handling as desired for structure and content of the integrated knowledge.

The next sections will show how a CA analysis is done in detail and discusses some examples of possible characteristics, which can be read from the results of the analysis. The analysis is done for objects and resources, which were already introduced. The examples include knowledge collections and containers. Example outputs of modules are given as is, meaning the figures show direct visualisation output from the implementation.

B. Comparison of data entities in and with collections

The following part illustrates characteristics of an integrated object and a knowledge resources collection object. Figure 12 shows the automatically computed CA module result for a case insensitive `vesuv ([Vv] [Ee] [Ss] [Uu] [Vv])` target for the above Gutenberg object instance (Figure 1) of the ordinary object [19].

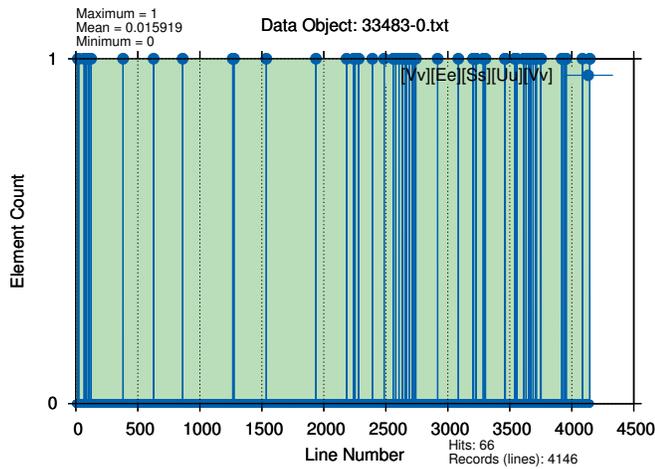


Figure 12. Comparative Analysis module result for a Gutenberg object precomputed by an assistance process for the case insensitive *vesuv* target.

The analysis including the illustration was automatically computed for the respective object. The results are shown on the background of a representation of the object entity. Resulting element counts displayed against line numbers reveal some respective characteristics of the object entity. The object entity extends over the maximum count / line number range (greenish color). Some assistive values are given, like number of records (lines in this case), number of hits, maximum, minimum, and mean value of hits. Despite the large number of hits, here, the mean value (blueish color) of hits is quite low, due to the relatively large number of records. The result also reveals sequences of higher hit-density and hit-patterns in the object entity in the illustration.

Figure 13 shows the automatically computed CA module result with the respective target (pattern) for the resulting knowledge resources collection object “Vesuvius” (Figure 3).

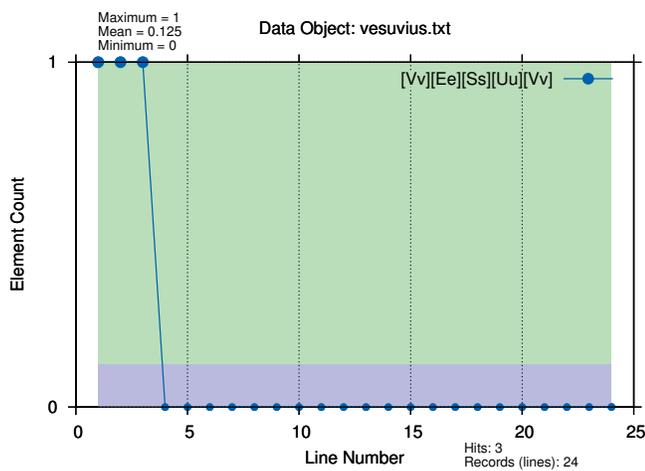


Figure 13. Comparative Analysis module result for the knowledge resources collection object “Vesuvius” (LX resources, geoscientific collection, excerpt).

The result shows some criteria of the object itself in context with the relevant mining pattern. The figure illustrates that the object contains a relevant mining result in the first and several

consecutive records (here: lines) with a maximal occurrence count of one in a record.

The density of relevant occurrences in the object is relatively high compared to common texts, even if from comparable special topic documents. Therefore, the mean value is quite high in that case. The computed background shading illustrates the space spanned by the available records (number of lines) and element counts. The mean value is illustrated by the border of the color change.

Both figures show that all terms were considered case insensitive. Here, collection objects and integrated object documents are chosen, which are referring to Vesuvius. For both, the mining workflow considers different ways of writing. Thus, for example, associations of Vesuvius and Pozzulan as well as directly linking to volcanic features and meteoric features via different resources.

The differences are intrinsic characteristics of the two types of objects, e.g., the length of the object, the concentration of hits, especially at the top of the object, and relatively high mean value of hits. The characteristics can be used to automatically decide to which extend objects can contribute to enhancements. Longer objects can contain several passages of records, which can be compared separately with an object and contribute to the enhancement of resources.

Features and properties of objects in collections are different to objects in containers. In many cases the differences of general objects in the Gutenberg resources can best be described in collections. Nevertheless, the comparison of collection and container objects will reveal characteristics of objects in both, which can contribute to the enhancement of workflow results.

C. Comparison of data entities in and with containers

Figure 14 shows the computed CA module result for a Gutenberg object instance (Figure 9, ordinary object [21]) for a case insensitive *vulc/volc* ([Vv][UuOo][Ll][Cc]) target.

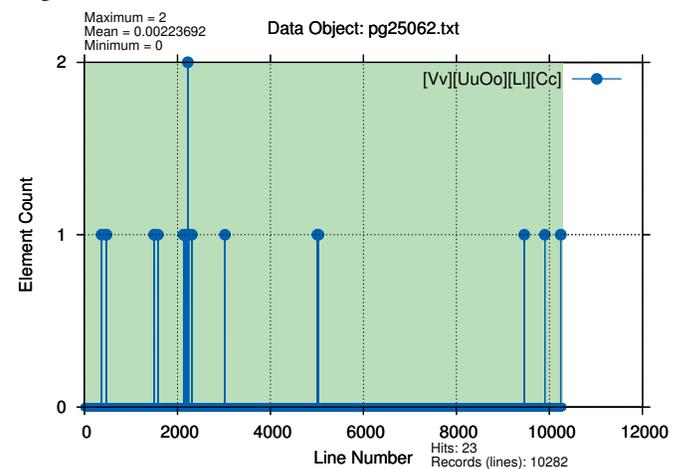


Figure 14. Comparative Analysis module result for a Gutenberg object precomputed by an assistance process for case insensitive *vulc/volc* target.

Figure 15 shows the automatically computed CA module result with the respective target (pattern) for the resulting knowledge resources collection object “Vesuvius” (Figure 3).

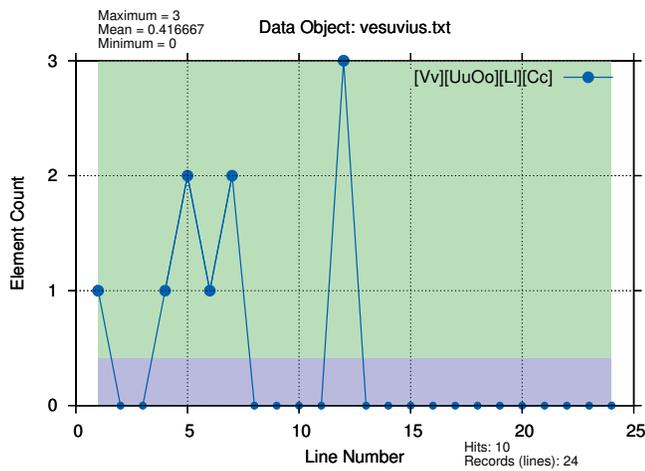


Figure 15. Comparative Analysis module result for the knowledge resources collection object “Vesuvius” (LX resources, geoscientific collection, excerpt).

There is more than one occurrence in several lines each, with a maximal occurrence count of three in a record. Figure 16 shows the computed CA module result for the volcanological features container (Figure 4) for the same target.

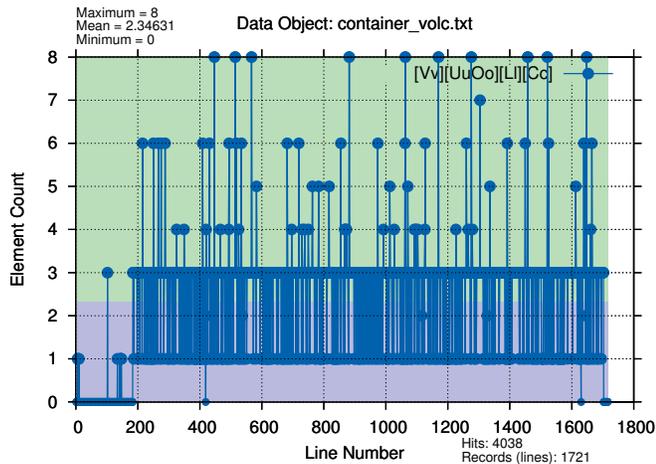


Figure 16. Comparative Analysis module result for the volcanological features container for case insensitive vulc/volc target (LX resources).

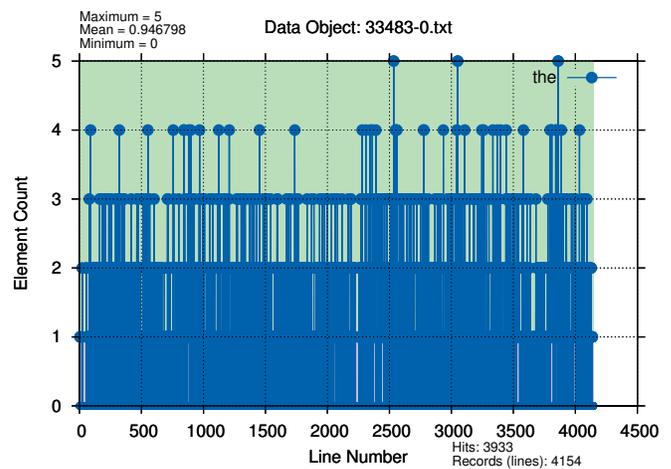
Both figures (Figures 16 and 14) illustrate the very high relevance of the objects. Nevertheless, the structure and density of hits is much higher in the container object than in the Gutenberg object. In addition, the mean value is extremely high for the container object. Also, the central part of the container object does not contain a line without the target. There are even more hits than records.

Even in a top hit Gutenberg object the number of records is much higher and the number of hits is lower. The comparison also reveals that both objects represent different object types, a knowledge resources object and a classical text object. The latter one mostly contains natural language. For any resources, many CA and Content Factor computations are done on a result matrix. With any workflow further information and decision making support can result from computing assistant views for the knowledge entities, e.g., based on their context.

D. Assistant views

For any type of resources, many CA and Content Factor computations are done on a result matrix. Besides the knowledge mining request “Vesuvius” in the context of “volcanology” with this example, the comparison with other requests can be helpful.

For example, common words, e.g., ‘the’, which does rarely occur in containers but mostly in book texts can be used for automated decision making. Figure 17 shows the computed CA module result for a Gutenberg object for a very common target, the.



knowledge resources objects may easily outnumber the number of Gutenberg documents. With the resources, the classification considers about fifty languages, summing up to about three million descriptions. Conceptual assistance is available for resource and object classification, which allows to automate integration workflows.

For the integration, instances of the objects containing the relevant data entities were automatically computed. It was possible to apply the provided means in the same way to the entities. The computation for the data entities from knowledge resources can be much more fine grained and systematic due to the complex structures and elements. The computation for the Gutenberg data entities can use the same means but some details and structure are not automatically available. The data sizes of the main Gutenberg data entities are most probably larger than these of the average knowledge resources' data entities.

The Content Factor method delivers an efficient tool in order to select object entities for their later integration with the knowledge resources and discovery workflows. The resulting objects can be easily CA computed and analysed regarding their contributions to enhancing the resources as well as the results of discovery.

Terms like 'precision' to not apply when dealing with complex knowledge, especially when knowledge entities are containing or are described by natural language content. In these terms, 'precision' depends on the question asked by the person implementing a scenario. Here, e.g., the CA mean values fill in for the purpose of a suitable measure. The range is from maximum fit to minimum fit.

B. CA mean values

Table V compares CA mean values from the computation for selected objects and target groups for the integrated resources.

TABLE V. SELECTED COMPUTATION DATA ENTITIES: OBJECTS AND TARGET GROUPS SORTED BY THEIR CA MEAN VALUES.

Object	Target / Target-Group	CA Mean
Knowledge res., Vesuvius	[Vv] [Ee] [Ss] [Uu] [Vv]	0.125
Gutenberg 33483-0	[Vv] [Ee] [Ss] [Uu] [Vv]	0.015919
Know. res., volc. feat. cont.	[Vv] [Ee] [Ss] [Uu] [Vv]	0.00291375
Gutenberg 25062	[Vv] [Ee] [Ss] [Uu] [Vv]	0.000486287
Know. res., volc. feat. cont.	[Vv] [UuOo] [Ll] [Cc]	2.34631
Knowledge res., Vesuvius	[Vv] [UuOo] [Ll] [Cc]	0.416667
Gutenberg 33483-0	[Vv] [UuOo] [Ll] [Cc]	0.0356971
Gutenberg 25062	[Vv] [UuOo] [Ll] [Cc]	0.00223692

There are entities with higher and lower mean values, for the Gutenberg resources as well as for the knowledge resources. Higher values indicate a cumulation of relevant terms, e.g., as with the appearance in collections, tabulars, and listings.

Practice showed that for complementing knowledge in the volcanological features container with extended context, relevant entities from the Gutenberg resources with higher mean values can be a primary source for references. Relevant entities from the Gutenberg resources with lower mean values may primarily deliver reference information for collection objects.

C. Comparisons in computation and analysis

As illustrated, longer objects can contain several passages of records, which can be compared separately with an object and contribute to the enhancement of resources. Beyond that, decisive workflows can benefit from the comparative results as to choose the best-fit objects for decisions, especially selecting the most fitting/associated passages in a text and comparing it to the most fitting/associated passage in another text.

Decisive workflows of that dimension are very challenging regarding computation and analysis. Table VI lists the counts of comparisons for the application in the above case study. The values are given for a single mining request (Figure 7).

TABLE VI. REPRESENTATIVE COUNTS OF COMPARISONS PER MINING REQUEST WITH THE CASE STUDY OBJECT ENTITIES.

Comparisons	Count
entities in collections	720,000
entities in containers	21,000
integrated documents, Gutenberg subset	570,000
overall, within filename/label space	550,000
overall, first result level	35,000
overall, second result level	15,000
overall, in pre-final results	4,500

A number of comparisons have to be done per mining request, depending on the complexity of the request and the configuration of the resources. In the above case, collections, containers, and integrated resources were used. Counts of comparisons were done in the respective entities and content, not in references outside the subsets.

In consequence, comparisons were done over all of the configured resources. The comparisons also included the filenames of the known integrated objects and the labels of the collection and containers entities because these regularly contain relevant knowledge for mining requests.

Most mining requests benefit from checking the result and trying to get additional information and enhanced results from the first results, using the same resources. In that case the first, second, and further result levels are considered intermediate results, which are used to go in depth and width for the consecutive mining.

The counts illustrate that a mining request can be very challenging, even for a subset of configured resources. There can also be needs to pre-cache and pre-compute resources, depending on the purpose of the application scenario in which the mining request should be embedded. In this case, the levels of the intermediate result generation can be considered an enhancement of selection, which leads to less comparisons required with the levels.

D. Ranking

For this scenario, a ranking was built from the rankings for the entities from the Gutenberg entities and from the knowledge resources.

The ranking considers the available information, e.g., classified targets, relevance of targets, references and context. The Gutenberg ranking especially considers the results from the CA, Content Factor and classification (LoC), based on the primary Gutenberg result matrix. The ranking of knowledge resources especially considers the CA, Content Factor, and classification, e.g., UDC and Universal Classified Classification (UCC). The integrated ranking considers the the CA, Content Factor values, and concordances of comparable entities.

An integration for a workflow ranking requires that the means need to be individually chosen for a certain application scenario. In this case, a records base (lines) was an appropriate choice for CA, Content Factor, and conceptual knowledge.

E. Computational trace and context

A common knowledge discovery process integrates a sequence of decision making processes at different levels, e.g., from which resources to which single objects. Each step in a sequence can require to handle millions of objects and references. The access to the Gutenberg resources is not intended to be automated. Therefore, no performance data is available for the Gutenberg resources or for conducting the precomputation for its whole content. The precomputation assistance includes the cached Gutenberg content for the respective mining targets.

Table VII shows the computation characteristics relevant with the workflow procedure for an example of the above integrated Gutenberg and knowledge resources case for two objects.

TABLE VII. COMPUTATION CHARACTERISTICS WITH THE WORKFLOW PROCEDURE FOR TWO INTEGRATED OBJECTS, WALL TIMES PER CPU.

<i>Workflow Procedure</i>	<i>Wall Time</i>
Precomputation assistance	24.8 s
Analysis, resources classification	1.2 s
Analysis, object classification	14.7 s
Comparative Analysis	3.2 s
(Integrative workflow step)	n s

The table times refer to one Central Processing Unit (CPU) per mining process (Intel Xeon, at 2.9 GHz). Due to the complexity of the elementary workflows it is not desirable to have more than one CPU per process involved at the atomic level. Arbitrary practical application scenarios involving many processes with large data resources may be organised to fit the architecture of the available infrastructure. With certain scenarios, where an author wants to integrate complex references, the precomputation assistance can benefit a lot from using many-CPU infrastructures.

The higher level workflow step, integrating the aforementioned procedures, will use a lot of intermediate results from procedures and content from resources. There is no general range for the time scale at the higher levels but at these levels the requirements on computation and communication can be extremely high, therefore, the higher level steps are candidates for parallelisation. Anyhow, workflow creators must

always be aware that computing requirements can be non-linear, depending on the workflow created by an author for a chosen purpose.

VIII. CONCLUSION

The paper presented the results of a research based on an advanced method for knowledge mining with multi-disciplinary knowledge resources and different data entities. Required modules and algorithms were successfully and efficiently implemented for supporting a Comparative Analysis, integrating different data entities in mining workflows.

This research showed that important characteristics of objects can be automatically identified and the results can be used for enhancing knowledge resources themselves as well as the discovery processes. The basis of the success is that with the availability of appropriate methodologies and methods different entities neither need to be left out from advanced knowledge mining workflows nor should their content be ignored. It was shown that in result, there is a complementary relationship between objects from knowledge resources and referred objects from external sources, including their data entities.

The Content Factor method for data description and analysis is used with all available resources. As was shown, CA methods cannot be replaced by other means like classification or Content Factor because they are based on completely different grounds but these complementary means can be integrated within more complex workflows. Classification is an integral complement in all parts of CA methods. In an implementation it means, e.g., that UDC can be used in all steps and components, starting from the knowledge resources. CA modules can help optimise the decision making, e.g., with supporting context-spanning Content Factor definition sets. CA modules can be used for delivering additional descriptive information, which can be used for documentation and knowledge mining purposes. CA is much beyond statistics. The significant part of the CA is the visualisation of pattern sequences in entities. The pattern sequences hold relevant parts of the entity characteristics and can also be used for documentation. The statistics are used in addition, for the analysis.

In consequence, the method allows an effective integration of resources and a dynamical selection of objects and entities. Selected objects and entities can be used dynamically, e.g., in discovery and decision making but also for persistently enhancing the context and references of available knowledge resources.

Objects from advanced knowledge resources can provide an excellent data base on knowledge. The knowledge resources can provide high quality object collections and containers with data entities of most reliable and unique content and qualities. Referred objects from external sources can extend the available data base regarding width and depth. Therefore, referred objects and external sources can extend the available data base and content. In the case study, the best fit targets regarding volcanological features from the resources, including the Gutenberg resources, were automatically analysed.

The method can enhance the quantity and quality of knowledge resources and discovery results by integration and analysis

of content and context for many application scenarios. In the case studies the quality was verified manually. A single measure of quality is out of scope of this research because the evaluation would depend on a certain purpose of application.

On the side of the Gutenberg resources, a number of challenges have been found especially with the Gutenberg objects themselves. With the documents, workflow creators face a lot of inconsistencies in structure and marking even regarding major elements. Bibliographic data and versioning could also be improved. Better structured and more complete bibliographic data would be beneficial for any wider and systematic use. A common container format for the Gutenberg documents, handling any data files and associated data in a flexible and 'clean' way would be beneficial.

The integration of the Content Factor method and the Comparative Analysis method is used in practice for progressive advancement of long-term multi-disciplinary knowledge resources and mining in a number of institutions and projects, including the knowledge resources used here for demonstration purposes. Special targets in practice are quality and quantity of entities but also balancing the content and context development of knowledge resources.

Information sciences, especially knowledge mining methods, have countless areas of implementation. Anyhow, the application scenarios of the methodology and the implemented methods are even not limited to knowledge mining. Besides the purpose laid out with this research, CA modules can be used as complementary and supportive methods applied with a wide range of advanced applications like document identification or plagiarism detection. For complex solutions, it can be desirable to integrate many components and modules. From the view of (the modern understanding of) technics, licensing acts as starting point for realisation of technical solutions. In this respect, licensing solutions are an integral part of almost any technical realisation and vice versa. Future work will be spent on further integrating different resources and creating methodologies, methods, and means for handling data entities and objects as well as on realisation aspects and education for the context of superordinate knowledge.

ACKNOWLEDGEMENTS

We are grateful to the "Knowledge in Motion" (KiM) long-term project, Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), for partially funding this implementation, case study, and publication under grants D2014F1P04518, D2014F2P04518, D2016F1P04683, and to its senior scientific members, especially to Dr. Friedrich Hülsmann, Gottfried Wilhelm Leibniz Bibliothek (GWLb) Hannover, Germany, to Dipl.-Biol. Birgit Gersbeck-Schierholz, Leibniz Universität Hannover, Germany, to Dipl.-Ing. Martin Hofmeister, Hannover, Germany, and to Olaf Lau, Hannover, Germany, for fruitful discussion, inspiration, practical multi-disciplinary case studies, and the analysis of advanced concepts. We are grateful to Dipl.-Ing. Hans-Günther Müller, Cray, Germany, for his work and assistance providing practical solutions to architectural challenges, practical private cloud and storage solutions, and excellent technical support. We are

grateful to all national and international partners in the Geo Exploration and Information cooperations for their constructive and trans-disciplinary support. We thank the Science and High Performance Supercomputing Centre (SHPC) for long-term support of collaborative research since 1997, including the GEXI developments and case studies.

REFERENCES

- [1] C.-P. Rückemann, "Comparative Analysis of Data Entities: Knowledge Mining Objects," in Proceedings of The Seventh International Conference on Advanced Communications and Computation (INFOCOMP 2017), June 25–29, 2017, Venice, Italy. XPS Press, Wilmington, Delaware, USA, 2017, Rückemann, C.-P., Flood, I. and Schwerdtfeger, I. and Simmendinger, C. and Beckett, G. (eds.), pages 17–23, ISSN: 2308-3484, ISBN-13: 978-1-61208-567-8, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2017_3_10_60019 [accessed: 2018-05-10].
- [2] C.-P. Rückemann, "Enhancement of Knowledge Resources and Discovery by Computation of Content Factors," in Proceedings of The Sixth International Conference on Advanced Communications and Computation (INFOCOMP 2016), May 22–26, 2016, Valencia, Spain. XPS Press, 2016, pages 24–31, ISSN: 2308-3484, ISBN-13: 978-1-61208-478-7, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2016_2_30_60047 [accessed: 2018-05-10].
- [3] C.-P. Rückemann, Z. Kovacheva, L. Schubert, I. Lishchuk, B. Gersbeck-Schierholz, and F. Hülsmann, Best Practice and Definitions of Data-centric and Big Data – Science, Society, Law, Industry, and Engineering. Post-Summit Results, Delegates' Summit: Best Practice and Definitions of Data-centric and Big Data – Science, Society, Law, Industry, and Engineering, Sep. 19, 2016, The Sixth Symp. on Advanced Computation and Information in Natural and Applied Sciences (SACINAS), The 14th Int. Conf. of Numerical Analysis and Applied Mathematics (ICNAAM), Sep. 19–25, 2016, Rhodes, Greece, 2016, URL: http://www.user.uni-hannover.de/cpr/x/publ/2016/delegatessummit2016/rueckemann_icnaam2016_summit_summary.pdf [accessed: 2018-05-10].
- [4] "LX-Project," 2018, URL: <http://www.user.uni-hannover.de/cpr/x/rprojs/en/#LX> [accessed: 2018-05-12].
- [5] "Project Gutenberg," 2018, URL: <http://www.gutenberg.org> [accessed: 2018-02-04].
- [6] Aristotle, Nicomachean Ethics, 2008, (Written 350 B.C.E.), Translated by W. D. Ross, Provided by The Internet Classics Archive, URL: <http://classics.mit.edu/Aristotle/nicomachaen.html> [accessed: 2018-05-10].
- [7] Aristotele, The Ethics of Aristotle, 2005, Project Gutenberg, eBook, eBook-No.: 8438, Release Date: July, 2005, Digitised Version of the Original Publication, Produced by Ted Garvin, David Widger, and the DP Team, Edition 10, URL: <http://www.gutenberg.org/ebooks/8438> [accessed: 2018-05-10].
- [8] Aristotele, Nicomachean Ethics, Volume 1, 2009, Project Gutenberg, eBook, eBook-No.: 28626, Release Date: April 27, 2009, Digitised Version of the Original Publication, Produced by Sophia Canoni, Book provided by Iason Konstantinidis, Translator: Kyriakos Zambas, URL: <http://www.gutenberg.org/ebooks/12699> [accessed: 2018-05-10].
- [9] T. Gooley, How to Read Nature: Awaken Your Senses to the Outdoors You've Never Noticed. New York, N.Y.: Experiment, 2017, ISBN: 978-1-61519-429-2.
- [10] L. W. Anderson and D. R. Krathwohl, Eds., A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Allyn & Bacon, Boston, MA (Pearson Education Group), USA, 2001, ISBN-13: 978-0801319037.
- [11] C.-P. Rückemann, "Creation of Objects and Concordances for Knowledge Processing and Advanced Computing," in Proceedings of The Fifth International Conference on Advanced Communications and Computation (INFOCOMP 2015), June 21–26, 2015, Brussels, Belgium. XPS

- Press, 2015, pp. 91–98, ISSN: 2308-3484, ISBN-13: 978-1-61208-416-9, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2015_4_30_60038 [accessed: 2018-05-10].
- [12] C.-P. Rückemann, “Advanced Association Processing and Computation Facilities for Geoscientific and Archaeological Knowledge Resources Components,” in Proceedings of The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2016), April 24 – 28, 2016, Venice, Italy. XPS Press, 2016, pages 69–75, ISSN: 2308-393X, ISBN-13: 978-1-61208-469-5, URL: http://www.thinkmind.org/download.php?articleid=geoprocessing_2016_4_20_30144 [accessed: 2018-05-10].
- [13] B. Gersbeck-Schierholz and C.-P. Rückemann, “Advanced References: Glue for Value Data,” KiM Sky Summit, Knowledge in Motion, September 18, 2016, Sky Summit Meeting, “Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF)”, Austria, 2016.
- [14] C.-P. Rückemann, “Advanced Knowledge Discovery and Computing based on Knowledge Resources, Concordances, and Classification,” International Journal On Advances in Intelligent Systems, vol. 9, no. 1&2, 2016, pp. 27–40, ISSN: 1942-2679, URL: http://www.thinkmind.org/download.php?articleid=intsys_v9_n12_2016_3 [accessed: 2018-05-10].
- [15] F. Hülsmann and C.-P. Rückemann, “Content and Factor in Practice: Revealing the Content-DNA,” KiM Summit, October 26, 2015, Knowledge in Motion, Hannover, Germany, 2015, Project Meeting Report.
- [16] F. Hülsmann, C.-P. Rückemann, M. Hofmeister, M. Lorenzen, O. Lau, and M. Tasche, “Application Scenarios for the Content Factor Method in Libraries, Natural Sciences and Archaeology, Statics, Architecture, Risk Coverage, Technology, and Material Sciences,” KiM Strategy Summit, March 17, 2016, Knowledge in Motion, Hannover, Germany, 2016.
- [17] “Library of Congress Classification Outline,” 2018, Library of Congress (LoC) Classification, URL: <https://www.loc.gov/catdir/cpsol/lcco/> [accessed: 2018-05-10].
- [18] M. Saderra Masó, Catalogue of Violent and Destructive Earthquakes in the Philippines, 2006, Project Gutenberg, eBook, EBook-No.: 18556, Release Date: June 11, 2006, Digitised Version of the Original Publication from 1910, URL: <http://www.gutenberg.org/ebooks/18556> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/18556/pg18556.txt> [accessed: 2018-02-04].
- [19] L. Palmieri, The Eruption of Vesuvius in 1872, 2010, Project Gutenberg, eBook, EBook-No.: 33483, Release Date: August 22, 2010, Digitised Version of the Original Publication from 1873, Translator: Mallet, Robert, (1810–1881), URL: <http://www.gutenberg.org/ebooks/33483> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/files/33483/33483-0.txt> [accessed: 2018-02-04].
- [20] M. Serao, Sterminator Vesevo (English: Vesuvius the great exterminator), 2014, Project Gutenberg, eBook, EBook-No.: 46658, Release Date: August 22, 2014, Digitised Version of the Original Publication from 1907, Diary of the Eruption of April 1906, URL: <http://www.gutenberg.org/ebooks/46658> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/46658/pg46658.txt> [accessed: 2018-02-04].
- [21] C. Davison, A Study of Recent Earthquakes, 2008, Project Gutenberg, eBook, EBook-No.: 25062, Release Date: April 12, 2008, Digitised Version of the Original Publication from 1905, URL: <http://www.gutenberg.org/ebooks/25062> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/25062/pg25062.txt> [accessed: 2018-02-04].
- [22] W. Hamilton, Observations on Mount Vesuvius, Mount Etna, and Other Volcanos, 2011, Project Gutenberg, eBook, EBook-No.: 35433, Release Date: March 1, 2011, Digitised Version of the Original Publication from 1774, Editor: Cadell, T., (1742–1802), URL: <http://www.gutenberg.org/ebooks/35433> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/35433/pg35433.txt> [accessed: 2018-02-04].
- [23] R. D’Awans, L’Ameublement de l’Hôtel de Pitsebourg au milieu du XVIIe siècle, 2004, Project Gutenberg, eBook, EBook-No.: 11586, Release Date: March 1, 2004, Digitised Version of the Original Publication from 1901, Communication faite en séance du 26 avril 1901, URL: <http://www.gutenberg.org/ebooks/11586> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/11586/pg11586.txt> [accessed: 2018-02-04].
- [24] A. H. C. Gelpke, Ueber die schrecklichen Wirkungen des Aufsturzes eines Kometen auf die Erde und über die vor fünftausend Jahren gehabte Erscheinung dieser Art, 2006, Project Gutenberg, eBook, EBook-No.: 18471, Release Date: May 29, 2006, Digitised Version of the Original Publication from 1835, URL: <http://www.gutenberg.org/ebooks/18471> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/18471/pg18471.txt> [accessed: 2018-02-04].
- [25] “QE: Science: Geology,” 2016, Library of Congress (LoC) Classification, URL: https://www.loc.gov/aba/cataloging/classification/lcco/lcco_q.pdf [accessed: 2018-05-10].
- [26] “DH: History: General and Eastern Hemisphere: Netherlands, Belgium, Luxemburg,” 2016, Library of Congress (LoC) Classification, URL: https://www.loc.gov/aba/cataloging/classification/lcco/lcco_d.pdf [accessed: 2018-05-10].
- [27] “QB: Science: Astronomy,” 2016, Library of Congress (LoC) Classification, URL: https://www.loc.gov/aba/cataloging/classification/lcco/lcco_q.pdf [accessed: 2018-05-12].
- [28] V. V. Vide, Sketches of Aboriginal Life, 2010, Project Gutenberg, eBook, EBook-No.: 33433, Release Date: August 14, 2010, Digitised Version of the Original Publication from 1846, URL: <http://www.gutenberg.org/ebooks/33433> [accessed: 2018-02-04], URL: <http://www.gutenberg.org/cache/epub/33433/pg33433.txt> [accessed: 2018-02-04].
- [29] “PS: Language and Literatures: American and Canadian literature,” 2016, Library of Congress (LoC) Classification, URL: https://www.loc.gov/aba/cataloging/classification/lcco/lcco_p.pdf [accessed: 2018-05-10].
- [30] “UDC 55: Earth Sciences. Geological sciences,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/032958> [accessed: 2018-05-10].
- [31] “UDC 9: GEOGRAPHY. BIOGRAPHY. HISTORY,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/068076> [accessed: 2018-05-10].
- [32] “UDC 93/94: History,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/068273> [accessed: 2018-05-10].
- [33] “UDC 94: General history,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/068284> [accessed: 2018-05-10].
- [34] “UDC 52: Astronomy. Astrophysics. Space research. Geodesy,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/027114> [accessed: 2018-05-10].
- [35] “UDC 821.111: English literature,” 2016, Universal Decimal Classification (UDC), URL: <http://udcdata.info/067893> [accessed: 2018-05-10].
- [36] “Multilingual Universal Decimal Classification Summary,” 2012, UDC Consortium, 2012, Web resource, v. 1.1. The Hague: UDC Consortium (UDCC Publication No. 088), URL: <http://www.udcc.org/udccsummary/php/index.php> [accessed: 2018-05-10].
- [37] “Creative Commons Attribution Share Alike 3.0 license,” 2012, URL: <http://creativecommons.org/licenses/by-sa/3.0/> [accessed: 2018-05-10].
- [38] C.-P. Rückemann, “Enabling Dynamical Use of Integrated Systems and Scientific Supercomputing Resources for Archaeological Information Systems,” in Proc. INFOCOMP 2012, Oct. 21–26, 2012, Venice, Italy, 2012, pp. 36–41, ISBN: 978-1-61208-226-4, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2012_3_10_10012 [accessed: 2018-05-10].
- [39] “UDC Online,” 2018, URL: <http://www.udc-hub.com/> [accessed: 2018-05-10].
- [40] “Universal Decimal Classification Consortium (UDCC),” 2018, URL: <http://www.udcc.org> [accessed: 2018-05-10].
- [41] F. Hülsmann and C.-P. Rückemann, “Summary on Algorithms and Workflows,” KiMrise, Knowledge in Motion Winter Meeting, December 12, 2014, Knowledge in Motion, Hannover, Germany, 2014.
- [42] “The Perl Programming Language,” 2018, URL: <https://www.perl.org/> [accessed: 2018-05-10].
- [43] “Gnuplot,” 2018, URL: <https://www.gnuplot.info/> [accessed: 2018-05-10].

Modeling of Gas Compressors and Hierarchical Reduction for Globally Convergent Stationary Network Solvers

Tanja Clees, Igor Nikitin, Lialia Nikitina, Łukasz Segiet
Fraunhofer Institute for Algorithms and Scientific Computing
Sankt Augustin, Germany

Email: {Tanja.Clees|Igor.Nikitin|Lialia.Nikitina|Lukasz.Segiet}@scai.fraunhofer.de

Abstract—Further development on globally convergent algorithms for solution of stationary network problems is presented. The algorithms make use of global non-degeneracy of Jacobi matrix of the system, composed of Kirchhoff’s flow conservation conditions and transport element equations. This property is achieved under certain monotonicity conditions on element equations and guarantees an existence of a unique solution of the problem as well as convergence to this solution from an arbitrary starting point. In application to gas transport networks, these algorithms are supported by a proper modeling of gas compressors, based on individually calibrated physical characteristics. This paper extends the modeling of compressors by hierarchical methods of topological reduction, combining the working diagrams for parallel and sequential connections of compressors. Estimations are also made for application of topological reduction methods beyond the compressor stations in generic network problems. Efficiency of the methods is tested by numerical experiments on realistic networks.

Keywords—modeling of complex systems; topological reduction; globally convergent solvers; applications; gas transport networks.

I. INTRODUCTION

This work is an extension of our conference paper [1], concentrated on mathematical modeling of single gas compressors in the context of designing globally convergent network solvers. Here we will add the modeling of aggregate compressor stations, obtained from the single compressors by parallel and sequential connections. We perform topological reduction of the stations, by combining the individually calibrated physical profiles of compressors into a cumulative profile, representing the compressor station as a single element. We will also estimate a benefit of generic topological reduction algorithms for generic networks and perform more numerical experiments for this purpose.

The simulation of transport networks in civil engineering has become increasingly important for the planning and stable operation of modern infrastructure. Compressors are essential elements in gas transport networks; they create pressure necessary for driving gas towards the consumers. A mathematical modeling of gas compressors should take into account their individually calibrated physical profiles. Our approach is based on conversion of the measured profiles into an explicitly resolved form suitable for globally convergent solvers. In particular, a proper signature of derivatives for the element equation of a compressor is provided.

Earlier [2], we have shown that the solvers for generic stationary network problems can be made globally convergent under special conditions on modeling of their elements. Stationary network problems combine linear Kirchhoff’s equations and (generally non-linear) element equations. The first

class of equations represents conservation laws, the second class describes the transport. We have proven that under certain monotonicity conditions on element equations, i.e., a special signature of the derivatives, the whole system possesses a globally non-degenerate Jacobi matrix. As a result, the problem always has exactly one solution. Moreover, standard algorithms, like Armijo backtracking line search and Katzenelson piecewise linear tracing, provide convergence to this solution from an arbitrary starting point.

These ideas have been implemented in our multi-physics network simulator MYNTS as described in [3], [4]. Considering gas transport networks, these papers used a simplified modeling of gas compressors, known in the simulation community as *free compressors*. This type of compressors does not possess limits on their power, only input or output pressure or gas flow are restricted. The present work extends the modeling by realistic characteristics to so called *advanced compressors*. Such compressors are described by individually measured physical profiles, defining the limits on power, revolution number, working region, etc. We will show how to incorporate such realistic characteristics into our globally convergent solver.

Modeling of gas transport networks has been considered in full detail in paper [5]. The networks are composed of a variety of elements (pipes, valves, compressors, drives, regulators, resistors, etc.), each type possessing a particular element equation. For instance, the pressure drop in pipes is described either by an empiric quadratic law [6] or by more accurate formulae by Nikuradse [7] and Colebrook-White [8]. Compressors and regulators have a control logic, implemented in the form of control equations or inequalities [5], e.g., a compressor/regulator can have a control goal to keep fixed output pressure, input pressure or flow value. There are various models for compressors (turbo, piston) and their drives (gas turbine, steam turbine, gas motor, electro motor), with characteristic diagrams calibrated on real engines.

The nodal variables (pressure, density, temperature, etc.) are related by equations of state, including various gas laws (ideal, Papay, standards of the American Gas Association – AGA and the International Organization for Standardization – ISO [9], [10]). Gas composition is defined in terms of molar components and effective gas properties (critical temperature and pressure, calorific value, molar mass, etc.), with appropriate equations describing their propagation and mixing. Thermal modeling [5] includes a number of physical effects (non-linear heat capacity, heat exchange with the soil, Joule-Thomson effect, i.e., a temperature drop due to free expansion of gas through a valve, regulator, etc.).

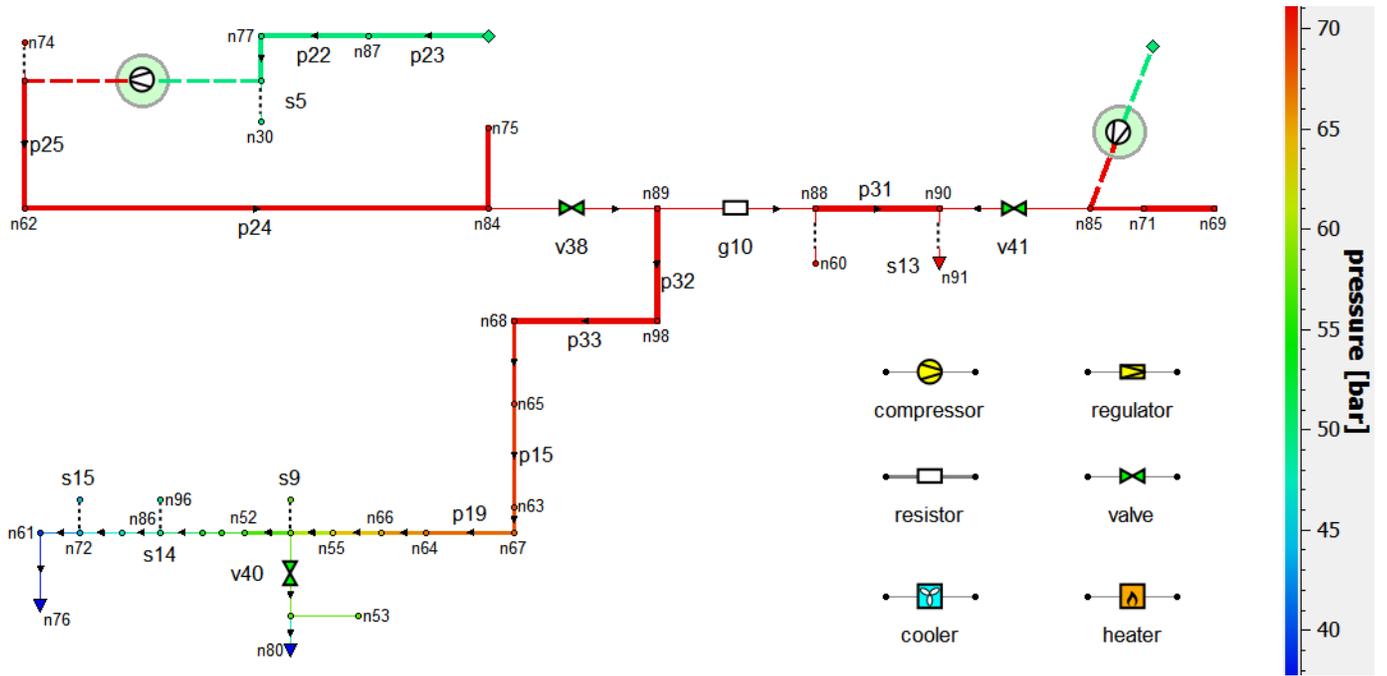


Figure 1. Test gas transport network N1 with 100 nodes, 111 edges.

The obtained system of equations and inequalities is solved by non-linear programming methods [11], [12]. Due to the non-linearity of equations, the stability of the solver critically depends on the choice of the starting point. For this purpose, various empirical strategies are used [13], [14]. An alternative has been proposed in our papers [2], [3], employing globally convergent algorithms, able to find the solution from an arbitrary starting point.

Topological reduction methods of the network utilizing the elimination of parallel and sequential connections are based on the concept of series-parallel graph (SPG, [15]). This is the graph reducible to a single edge by repeated application of such operations. Further extension is a generalized series-parallel graph (GSPG, [16]), where in addition to the elimination of parallel and sequential connections one can eliminate a leaf (node of valency 1). SPGs and GSPGs are recognizable in linear time with respect to the size of the graph, or in logarithmic time using a linear number of parallel processors. Decomposition of such graphs to elementary components is also performed in linear time, as well as solution of many other graph-theoretical problems, which would be NP-complete for generic graphs [17–19]. In our application, the concepts of SPGs and GSPGs are of key importance, since they allow essential reduction of the networks to an irreducible skeleton with small number of elements. The solver should be applied only to the skeleton, while the complete solution can be reconstructed with simple algorithms using the reduction history. The efficiency of such reduction for generic networks will be studied in this paper.

In Section II, we recall conditions on the generic stationary network problem, necessary for global convergence, and concretize these conditions in application to gas transport networks. In Section III, we describe modeling of advanced gas compressors. In Section IV, we present our implementation of

modeling, which fulfills the conditions for global convergence. In Section V, we describe topological reduction algorithms for gas compressor stations and in Section VI generalize them to arbitrary elements and network types. In Section VII, we present numerical experiments with a number of realistic gas transport network examples and discuss the results obtained.

II. GLOBAL CONVERGENCE AND GAS TRANSPORT NETWORKS

A generic stationary network problem can be written as

$$\sum_e I_{ne} Q_e = Q_n^{(s)}, \quad n \notin N_P, \quad P_n = P_n^{(s)}, \quad n \in N_P, \\ f_e(P_{in}, P_{out}, Q_e) = 0, \quad (1)$$

where indices $n = 1 \dots N$ denote the nodes and $e = 1 \dots E$ the edges of the associated network graph, I_{ne} is an incidence matrix of the graph, Q_e are flows through the edges, $Q_n^{(s)}$ are source/sink contributions, localized in supply/exit nodes, P_n are nodal variables (pressure for gas transport networks), $P_n^{(s)}$ are set values, localized in the subset N_P of supply/exit nodes, at least one value per connected component of the graph. Let the element equations possess derivatives of the signature:

$$\partial f_e / \partial P_{in} > 0, \quad \partial f_e / \partial P_{out} < 0, \quad \partial f_e / \partial Q_e < 0. \quad (2)$$

It has been proven in [2] that the system (1) under condition (2) possesses a globally non-degenerate Jacobi matrix.

Gas transport networks, e.g., the networks shown in Figure 1, consist of several types of elements, all possessing the property (2). The gas networks are mostly composed of pipes with a non-linear (nearly quadratic) element equation. Some elements (valves and shortcuts) have linear equations, most complex elements (compressors and regulators) possess piecewise linear equations. According to [14], all continuous

piecewise linear functions can be represented in a max-min form:

$$f(x) = \max_i \min_j \sum_k a_{ijk} x_k + b_{ij}, \quad (3)$$

where a, b are coefficient lists. In particular, free compressors are described by the following element equation:

$$\begin{aligned} \max(\min(P_{in} - P_L, -P_{out} + P_H, -Q + Q_H), \\ P_{in} - P_{out}, -Q) + \epsilon(P_{in} - P_{out} - Q) = 0. \end{aligned} \quad (4)$$

The compressor tries to satisfy one of the following control goals:

- a specified pressure on output (SPO);
- a specified pressure on input (SPI);
- a specified mass flow (SM).

Being combined with the given upper and lower bounds:

$$\begin{aligned} P_H &= \min(SPO, POMAX), \\ P_L &= \max(SPI, PIMIN), \\ Q_H &= \min(SM, MMAX), \end{aligned} \quad (5)$$

the element equation defines a polyhedral surface shown in Figure 2, top. Here, P_H stands for high pressure limit, P_L – low pressure limit, Q_H – high flow limit; $POMAX$ is an upper safety bound on output pressure, $PIMIN$ is a lower safety bound on input pressure, $MMAX$ is an upper safety bound on the flow.

Every face of the diagram corresponds to the best possible fulfillment of the control goal, e.g., $P_{out} = P_H$ (typical for SPO-mode), $Q = Q_H$ (typical for SM-mode), $P_{in} = P_{out}$ (bypass BP, equivalent to an open valve), $Q = 0$ (OFF, equivalent to a closed valve), etc. A small ϵ value is used for regularization purposes.

Every compressor is part of a compressor station, the simplest one is shown in Figure 2, bottom. This figure presents a compressor station with one machine unit, including (in flow direction) input resistor, compressor, cooler, output resistor, exit valve, (in an oblique direction) bypass valve, bypass regulator, both currently closed. In more complex stations, the compressors with accompanying elements are cascaded to parallel or sequential configurations, see Figure 3.

III. ADVANCED MODELING OF GAS COMPRESSORS

In addition to pressure P , the nodes in gas transport networks possess other common variables, including: ρ – mass density, T – temperature, z – compressibility factor, μ – molar mass.

In addition to conserving mass flow Q , measured in kg/s, sometimes volume flow Q_x is also considered. It is measured in m³/s with explicit reference to the measurement conditions, e.g., Q_{norm} represents the volume flow under normal conditions (1 bar, 273.15 K), $Q_{in,out}$ refers to the volume flow under conditions in the input and output nodes. Different flow definitions are related by the formula:

$$Q = Q_{norm} \rho_{norm} = Q_{in} \rho_{in} = Q_{out} \rho_{out}. \quad (6)$$

Advanced compressors bring four new variables: H_{ad} – adiabatic enthalpy increase, η_{ad} – adiabatic efficiency, r – revolution number of compressor drive, W – power of compressor

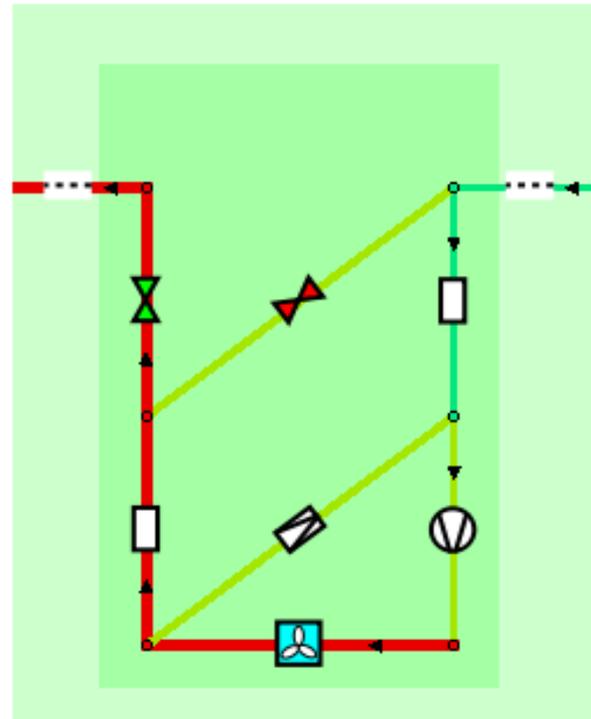
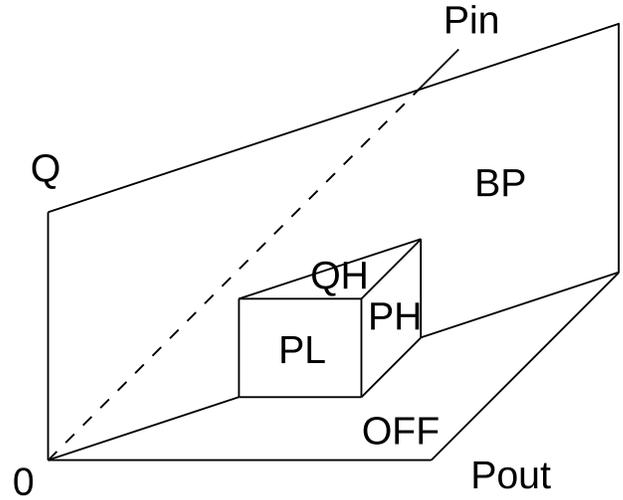


Figure 2. On the top: control diagram of free compressor. On the bottom: a compressor station with one machine unit.

drive. Also, four equations are added [5]:

$$\begin{aligned} H_{ad} &= RT_{in} z_{in} / (\mu \alpha) \cdot ((P_{out}/P_{in})^\alpha - 1), \\ W &= Q H_{ad} / \eta_{ad}, \\ H_{ad} &= (1, r, r^2) \cdot A \cdot (1, Q_{in}, Q_{in}^2)^T, \\ \eta_{ad} &= (1, r, r^2) \cdot B \cdot (1, Q_{in}, Q_{in}^2)^T, \end{aligned} \quad (7)$$

where R is the universal gas constant, κ the adiabatic exponent, $\alpha = (\kappa - 1)/\kappa$. A and B are (3x3)-matrices filled with calibration constants. In addition, working limits for compressors

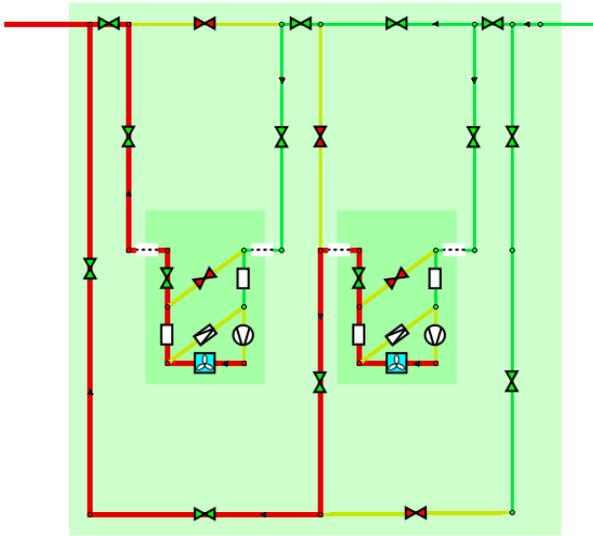


Figure 3. A compressor station with two machine units in parallel configuration.

are defined by the following inequalities:

$$\begin{aligned} r_{min} &\leq r \leq r_{max}, \quad \eta_{ad} \geq \eta_{min}, \\ Q_{in} &\geq Q_{min}, \quad W \leq W_{max}, \\ Q_{min} &= (1, H_{ad}, H_{ad}^2) \cdot C^T, \\ W_{max} &= (1, r, r^2) \cdot D^T, \end{aligned} \quad (8)$$

where the constants r_{min}, r_{max} define limits of the revolution number, η_{min} – the lower limit on efficiency (so called choke line); Q_{min} is a lower limit on the input volume flow (a surge line), W_{max} is an upper limit on power. C, D are (3)-vectors filled with calibration constants.

The constants in (A, B, C, D) are found by fitting the measured data for the compressor considered as part of a calibration procedure and further represent the individual profiles for this compressor. Here, we described the modeling for a common class of turbo compressors and gas turbine drive engines. The other types are simpler in implementation and can be modeled analogously.

Figure 4 (top) shows profiles for a typical turbo compressor. In this plot, the horizontal axis represents input volume flow Q_{in} , the vertical axis – adiabatic enthalpy increase H_{ad} . Solid blue curves are the lines of constant revolution number r , their uppermost curve corresponds to r_{max} , the lowest curve – to r_{min} . The red curve is the surge line $Q_{in} = Q_{min}$, while the rightmost cyan curve – the choke line $\eta_{ad} = \eta_{min}$. The points in this diagram depict the data measured, a blue cross denotes the current working point of the compressor.

The equations (7) serve as definitions of newly introduced variables, while the inequalities (8) define the restrictions, in addition to (4) of those for a free compressor. The upper bounds $r = r_{max}$ and $W = W_{max}$ define new upper bounds for the flow and should be combined with the one defined by the diagram for a free compressor, shown in Figure 2 top. The lower bounds $r = r_{min}$ and $Q_{in} = Q_{min}$ show the points where the station automatically opens its bypass

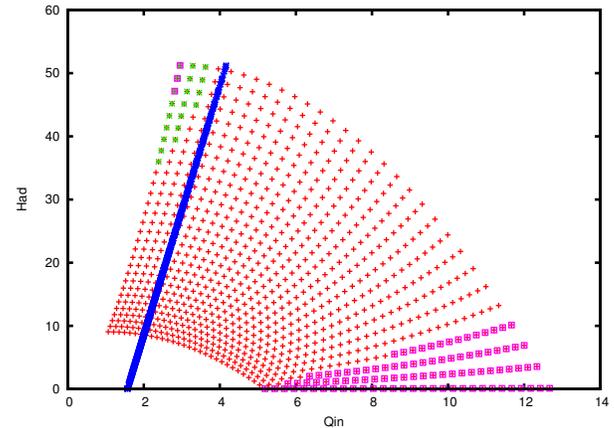
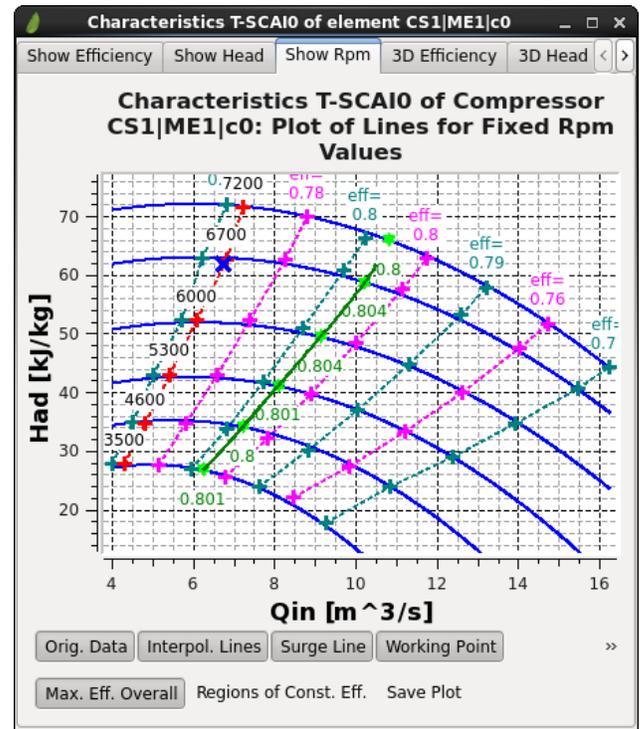


Figure 4. Advanced modeling of compressors. On the top: typical characteristic diagram. Solid blue curves are lines of constant revolution number. On the bottom: stability analysis.

regulator, shown in Figure 2 bottom. After that the gas starts to circulate inside the station, so that the compressor never violates its bounds. The choke line $\eta_{ad} = \eta_{min}$ cuts off a region of unstable calibration related with the small η_{ad} in the denominator of (7). Usually, the working point of a compressor is not located in this region, except of the starting procedure. On necessity the diagram can be continued in this region by a convenient monotone formula.

If P_{in} and P_{out} are fixed and the compressor is on its $r = r_{max}$ limit, it is straightforward to resolve the equations analytically, finding $H_{ad}, Q_{in}, \rho_{in}, Q, \eta_{ad}$ and W , in this order. If r and Q_{in} are fixed and the compressor is on its $W = W_{max}$ limit, the equations can be resolved in

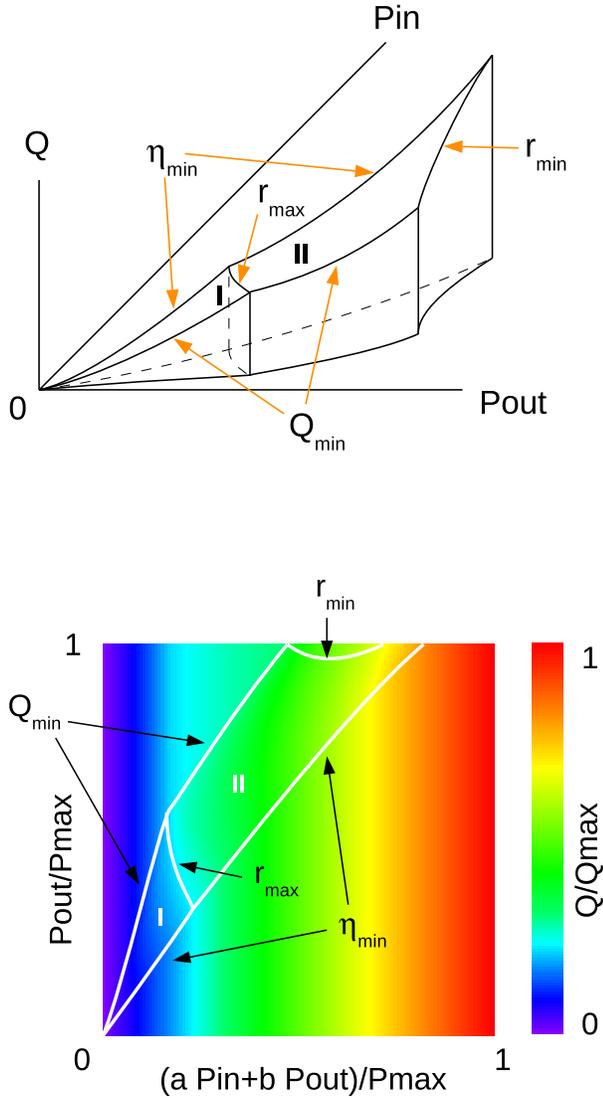


Figure 5. Advanced modeling of compressors (cont'd). On the top: control diagram of an advanced compressor. On the bottom: the same diagram in affine coordinates and color map representation.

the order H_{ad} , η_{ad} , W_{max} , Q , ρ_{in} , P_{in} , P_{out} . The first analytic formula gives an explicit representation $Q(P_{in}, P_{out})$ for the surface, defining a patch of the element equation. The second one represents the other patch in a parametric form $(r, Q_{in}) \rightarrow (P_{in}, P_{out}, Q)$. By numerical differentiation, it is possible to find the normals to both surfaces, which directly define the signatures of the corresponding patches. We recall that the correct signature reads $(+ - -)$, see (2). In Figure 4 (bottom) green points show the area of wrong signature. The blue curve is the surge line, magenta points show the area $\eta_{ad} < \eta_{min}$. Normally, the unstable green area is cut off by the surge and choke lines, so that the whole working region of the compressor is stable. In rare cases, when it is not so, the surge and choke lines should be modified accordingly.

IV. GLOBALLY CONVERGENT IMPLEMENTATION

Figure 5 (top) shows the surface $Q_{adv}(P_{in}, P_{out})$ defined by the characteristics of the advanced compressor. The cali-

brated part of the surface is located between the curves Q_{min} and η_{min} and consists of two patches I and II, connected on the r_{max} line. Patch I is located between the r_{max} line and the origin. Here, the input and output pressures are small and the compressor's performance is limited by its maximal revolution number: $r = r_{max}$. Patch II is located between the r_{max} and r_{min} lines. Here, the pressures are large and the compressor is limited by the maximal power of the drive: $W = W_{max}$. On Q_{min} and r_{min} curves the surface vertically falls down. This behavior corresponds to the open bypass regulator. Q in that case denotes a total mass flow through the compressor and bypass regulator. The flow through the compressor remains equal to Q_{min} or the equivalent flow on the r_{min} line, while the negative difference ΔQ circulates through the bypass regulator. A slope of the vertically falling faces should be ϵ -regularized to provide the necessary signature $(+ - -)$. The surface should be continued beyond η_{min} curve by any function supporting the same signature.

Figure 5 (bottom) shows the same diagram as a color map. Patch II requires a conversion from parametric to explicit representation. For this purpose, we adopt resampling algorithms well known in computer graphics (CG). At first, we perform an affine transformation:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_{in} \\ P_{out} \end{pmatrix} \cdot \frac{1}{P_{max}}, \quad (9)$$

with $a + b = 1$. The square on the (x, y) -plane is represented as a $N_{px} \times N_{py}$ pixel buffer, storing floating point values of Q in double precision. Patches I and II are regularly sampled and represented as triangle strip sets. Then, the patches are rendered onto the (x, y) -plane using the Z-buffer algorithm. Finally, the remaining gaps are filled by copying a constant Q -value along the columns and to the right – by a linearly increasing function in the row. As a result, the Q_{adv} -function in these regions becomes dependent only on x . Monotonous increase of Q on the border lines and a choice of affine coefficients $a > 0$, $b < 0$ support correct signature of the function $Q_{adv}(P_{in}, P_{out})$. If the bypass regulator is activated, the part above the upper border lines must be reset towards $Q = 0$, providing a regularized vertical fall of the surface on this bound.

The described algorithms provide a transformation from the calibration coefficients (A, B, C, D) and characteristic diagram in Figure 4 (top) to the tabulated function $Q_{adv}(P_{in}, P_{out})$, represented by the color map in Figure 5 (bottom). This transformation should be done once per advanced compressor. For a moment, we use an implementation of the CG algorithms on Central Processing Unit (CPU) and plan their acceleration with Graphics Processing Unit (GPU).

In the solver, the lookup function $Q_{adv}(P_{in}, P_{out})$ is made available via rapid bilinear interpolation of tabulated values inside the (x, y) -square. It is continued to the whole (x, y) -plane by an explicit analytic formula:

$$f(x, y) = f(\hat{x}, \hat{y}) + k(\min(x, 0) + \max(x - 1, 0)), \quad (10)$$

$$\hat{x} = \min(\max(x, 0), 1), \quad \hat{y} = \min(\max(y, 0), 1),$$

with a constant $k > 0$. This global function is constructed similarly to the continuation formulas in [2]. In our special case it provides monotonous increase in x and constancy in y outside of the tabulated region.

Finally, the element equation for an advanced compressor is obtained by extending (4) as follows:

$$\max(\min(P_{in} - P_L, -P_{out} + P_H, -Q + Q_H, \underline{Q_{adv}(P_{in}, P_{out}) - Q}, P_{in} - P_{out}, -Q) + \epsilon(P_{in} - P_{out} - Q) = 0. \quad (11)$$

For clarity, the inserted term is underlined.

We have implemented the algorithms described above in our network simulator MYNTS in a preliminary version (solver strategy “stable”).

V. MODELING OF GAS COMPRESSOR STATIONS

Machine units consisting of compressors and their drives are often combined in stations using parallel and sequential connections, such as shown in Figure 6. Parallel connections (Figure 6a,c) are used to increase throughput, while sequential connections (Figure 6b,d) allow to increase the output pressure in several steps. Mixed connections are also used, possessing various topologies (Figure 6e-g). In these figures, the compressor symbols represent single compressor stations (Figure 2), where the functional element is the compressor and the other elements serve mainly for flow switching.

The idea considered in this section is to combine the element equations of single compressors and to represent the station as a single element. At first, free diagrams depending on (SPO,SPI,SM) parameters should be unified with the advanced characteristics, so that the algorithms can operate on a single table per compressor. Then, the corresponding combination of lookup tables will describe the behavior of the station as a whole. This will reduce the number of equations in the system and will also allow to resolve certain stability issues.

Compressor stations bring unstable internal degrees of freedom, i.e., undefined balance of flows in parallel stations, undefined intermediate pressure in sequential stations. This instability is a particular case of the common singularity, for which ϵ -regularization is used in the element equation (11). From a geometrical point of view, the hypersurfaces defined by the equations do not have a stable transversal intersection, but are in unstable tangent position.

The reason of this instability is the appearance of linear dependencies in the system matrix. For instance, two parallel SPO compressors set the same pressure value in the same output node, imposing two identical $P = SPO$ equations in the system. The system matrix becomes formally degenerate. Also, the fact that one equation is actually wasted leads to the effective reduction of the system size by one, so that one of the variables remains undefined. In this particular example, the undefined variable is the flow balance between the compressors. In the other case, namely a sequential connection of SM compressors, the $Q = SM$ equation is duplicate and the undefined variable is the intermediate pressure between the compressors.

The replacement of the station by a single element with combined characteristics removes internal degrees of freedom and improves stability of the numerical simulation.

Compressors, as well as all other elements are described by functions of the form: $Q(P_{in}, P_{out})$, or equivalently: $P_{out}(P_{in}, Q)$, $P_{in}(P_{out}, Q)$, where Q is the mass flow, P_{in}, P_{out} are input/output pressures. Monotonicity conditions

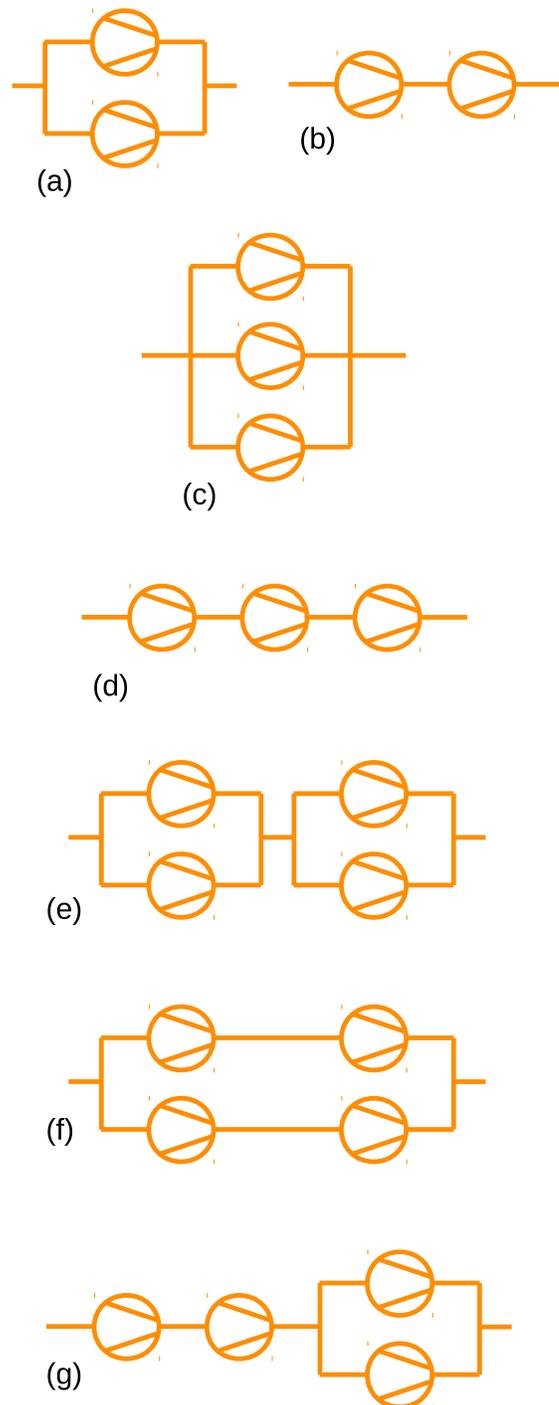


Figure 6. Examples of compressor stations: (a) parallel, (b) sequential, (c) 3 parallel, (d) 3 sequential, (e-g) mixed.

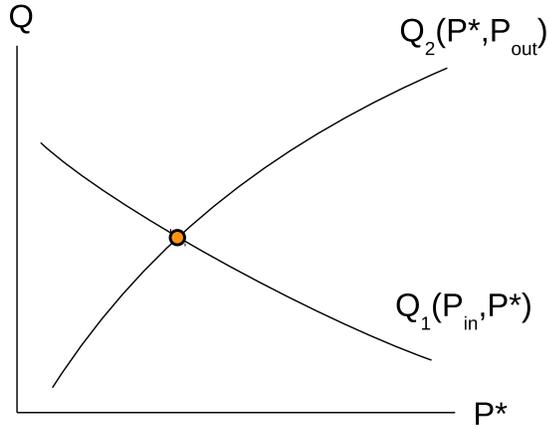


Figure 7. Equality of flows in sequential station has a single solution, due to monotonicity of element equations.

allow a unique inversion for any of these functions with respect to one of the arguments, passing from one representation to the other.

According to Kirchhoff's law, for connected elements (1, 2) these functions can be combined as follows:

parallel connection:

$$Q(P_{in}, P_{out}) = Q_1(P_{in}, P_{out}) + Q_2(P_{in}, P_{out}); \quad (12)$$

sequential connection:

$$P_{out}(P_{in}, Q) = P_{out,2}(P_{out,1}(P_{in}, Q), Q). \quad (13)$$

Considering the monotonicity of the element equations, it is easy to verify that the combined elements possess correct signature of the derivatives, necessary for the global convergence. Indeed, the sum of monotonic functions in parallel connection is the monotonic function of the same signature. The composition of functions in sequential connection is monotonously increasing with P_{in} and monotonously decreasing with Q , as needed.

Practically, if the elements are represented by a lookup function $Q(P_{in}, P_{out})$, their parallel connection becomes especially simple, one just needs to sum the corresponding lookup tables. Sequential connection requires the inversion of the lookup functions, which can be performed on the fly by solving the equation

$$Q_1(P_{in}, P^*) = Q_2(P^*, P_{out}) \quad (14)$$

for the intermediate pressure P^* . Here the first function monotonously decreases with P^* , while the second function monotonously increases with P^* . Also, the continuation of the functions outside working region ensures the following signs at infinities

$$Q_1(P_{in}, \pm\infty) = \mp\infty, \quad Q_2(\pm\infty, P_{out}) = \pm\infty. \quad (15)$$

As a result, the graphs of these functions always have a single intersection point, as shown in Figure 7. Thus, equation (14) has a unique solution. It can be found, e.g., by Newton's method with Armijo line search stabilizer. Since lookup functions are piecewise linear with respect to P^* , this method converges to the solution in a finite number of steps. Then, the

resulting $Q_1(P_{in}, P^*)$ gives a combined diagram $Q(P_{in}, P_{out})$ for the sequential station.

Being equipped with the elementary algorithms for parallel and sequential connections, one can apply them recursively to reduce the compressor station to a single element, also for mixed configurations.

VI. GENERIC TOPOLOGICAL REDUCTION

In exactly the same way, the rules for combining parallel and sequential connections can be applied to the edge elements of an arbitrary type. Suppose that all elements in the network are represented by tabulated lookup functions. Applying the above described parallel and sequential connection rules, one can combine the elements and reduce the network graph significantly. In the given section, we will investigate this possibility. Here, we do not actually implement the tabulation for all elements, instead, we estimate the efficiency of the reduction to forecast a benefit of such implementation.

Parallel and sequential connection rules are based on Kirchhoff's law, which must be modified in supply/exit nodes. The simple summing rule for parallel configuration remains valid also in the presence of such nodes. For sequential configuration modifications of the algorithm are needed.

There are Q_{set} nodes, where source/sink contributions for the flow are given. If an intermediate node in sequential configuration has Q_{set} , it should be used to shift one of the flow arguments in $P_{out,2}(P_{out,1}(P_{in}, Q), Q - Q_{set})$, or, equivalently, in the equation $Q_1(P_{in}, P^*) - Q_{set} = Q_2(P^*, P_{out})$. The shifts $Q - Q_{set}$ obviously do not change the signature of the derivatives in the element equation and do not influence the stability.

To keep a standard form of the element with a unique flow assigned to it one can formally define $Q = Q_1$ and move Q_{set} to the output node. Such shifts should be stored so that the inverse reconstruction can return all Q_{set} entries to their places. The algorithm will shift Q_{set} entries until they stop in a node of irreducible skeleton. If two Q_{set} entries collide, their values can be summed. If Q_{set} stops in P_{set} node, it means an addition to the unknown Q source/sink contribution, which is already present there.

In P_{set} nodes Kirchhoff's law is replaced by $P = P_{set}$ condition. If this node appears as the intermediate node in sequential configuration, the element functions $Q_1(P_{in}, P_{set})$ and $Q_2(P_{set}, P_{out})$ can be used to find (generally non-equal) flows. The difference of these flows is sourced/sinked in this P_{set} node.

However, to keep the element standard, one needs to shift the unknown flow contribution to the output node. It cannot be implemented as a standard Q_{set} since this contribution is not a constant but depends on P_{in}, P_{out} . Therefore, the topological reduction should be stopped in P_{set} nodes.

Practically, the algorithm can be also stopped in the other user-defined nodes. One can agglomerate passive sub-graphs, e.g., consisting of pipes and resistors, and keep P_{set}, Q_{set} , compressors, regulators and other active elements non-contracted. In this way, a significant reduction can be achieved while the state of active elements and all values defining scenario settings remain under user control.

In particular cases, further simplifications are possible. Electric elements depend on voltage difference and require not

2D but 1D profiles. This significantly reduces the computational effort. Linear electric resistors and quadratic hydraulic resistors are described by numerical constants instead of profiles, bringing an additional speedup.

Our non-linear element equations are similar to Green's functions [20] in linear systems measuring a response of the system on an external perturbation. The elements with two ends (edges) are described by two-point Green's functions, which can be properly combined into the other two-point Green's function by parallel and sequential connections. For more complex topologies, higher order response functions can be introduced (three-point, etc.). Technically it is possible to precompute also higher order functions. However, their storage and processing will require introduction of multidimensional tables with enormous consumption of resources and computational expenses. We prefer to stay on two-point level and apply the reduction algorithms there.

It is clear that not all networks can be reduced to a single element by a combination of parallel and sequential (parseq) reduction algorithms. Figure 8 shows some examples of irreducible network topologies. Circular symbols on this figure denote arbitrary, not necessarily identical elements. In the examples shown in Figure 8a-d,g neither parallel nor sequential connections can be found and the networks cannot be further reduced by these algorithms. An exception is the case when the vertical elements in the configurations Figure 8a,b are closed or open valves, or shortcuts. Then these configurations become further reducible.

An additional resource for the reduction is available for tree-like subgraphs, as shown in Figure 8c,d. Suppose that the leaves of the tree (nodes of valency 1) are all Q_{set} nodes, including the regular $Q_{set} = 0$ ones. An exception is one node (the root), which is either P_{set} node or it is connected to a larger parent graph. For definiteness, on Figure 8c,d, the leftmost node can be considered as the root.

Using Kirchhoff's law, one can sum the flows starting from leaves towards the root and completely determine the flow distribution in this subgraph. Then, starting from the root and using element functions in the form $P_{out}(P_{in}, Q)$, one can transfer P -values upto the leaves. In this way, the problem can be solved in the subgraph by a simple algorithm without actually passing it to the generic solver. On the topological level, the whole tree-like subgraph can be collapsed to the root node, representing a cumulative Q_{set} for this subgraph, in frames of the parent graph.

The tree reduction algorithm is essentially based on the presence of a single P_{set} node in the subgraph. If the tree contains two or more P_{set} nodes, Kirchhoff's law cannot be applied and the reduction algorithm stops there. The reduction can also require the combination of all algorithms. On Figure 8e one needs to apply parallel reduction to proceed with the tree. On Figure 8f the graph can be reduced by application of sequential, parallel and tree algorithm. On Figure 8g one needs to apply tree algorithm first, then parseq.

It is remarkable that namely the operations of parallel/sequential elimination and tree reduction, which in our application simplify the numerical solution of network problems, are used in graph theory to accelerate the solution of combinatorial problems, upto linear time performance. In our case, the elimination of subgraphs of SPG/GSPG type [15],

[16] allows to reduce the network to a skeleton, on which the actual numerical solution should be performed. Then the solution in the reduced parts can be reconstructed back by simple algorithms.

In our implementation, the topological reduction procedure starts from a common algorithm (clean), performing the following steps. All closed elements (closed valves, compressors, regulators) are removed from the network together with disconnected subgraphs where no P_{set} is reachable. This step ensures that P -values are everywhere defined. All open elements (open valves, shortcuts, short segments of pipes, compressors and regulators in bypass mode, etc.) that transfer P -values from one node to the other without a change are collapsed. This step ensures that no superconductive loops appear leading to cycles of undefined flow. The both steps are necessary to provide global non-degeneracy of the system.

Then parallel, sequential and tree algorithms are applied one after the other, repeatedly, till the reduction is complete. The reduced graph is passed to the solver. After the solution is found, the reduction history can be used to recover the complete answer.

We remind that the last step in the reduction procedure was either a join of two elements, parallel or sequential, or elimination of the leaf. If it was parallel join, P_{in}, P_{out} are common, Q is known, while Q_1/Q_2 balance can be found using stored $Q_1(P_{in}, P_{out}), Q_2(P_{in}, P_{out})$ characteristics. If it was sequential join, Q is common, P_{in}, P_{out} are known, the intermediate value P^* can be reconstructed using stored $P_{out,2}(Q, P_{in})$. If it was elimination of the leaf, Q and P_{in} are known, P_{out} can be found from the stored $P_{out}(Q, P_{in})$. Thus, we have recovered the missed properties in one step and can repeat this process recursively until all these data are reconstructed.

VII. RESULTS

For benchmarking the algorithms, we have received a number of realistic test scenarios from our industrial partners. The simplest gas transport network N1 from our test set is shown in Figure 1. It contains two compressor stations, each equipped with two machine units, cf. Figure 2 (bottom), working in parallel mode. It has two P_{set} supplies and three Q_{set} consumers. The color shows the pressure distribution over the network, arrows – the direction of gas flow, thickness of lines – the diameter of the pipes. Supplies are shown by rhombi, consumers (n76, n80, n91) are shown by triangles. The main elements are shown in the legend.

A closeup to one of the compressors is shown in Figure 2 (bottom), and its characteristics are displayed in Figure 4 (top). Parameters of more complex networks are presented in Table I. In particular, medium-sized network N2 contains about a thousand nodes and edges and is equipped with 7 compressors. The largest considered network N3 has about five thousand nodes and edges and is driven by 25 compressors. Topological connection of elements in the network together with geographic coordinates, the lengths and diameters of pipes form so called *geometry* of the network. Physical setting, such as supply pressures and consumer flows, control settings of compressors and regulators, define *scenario* for the particular simulation case. All networks in our test set were simulated with the same gas composition at 20°C environmental temperature.

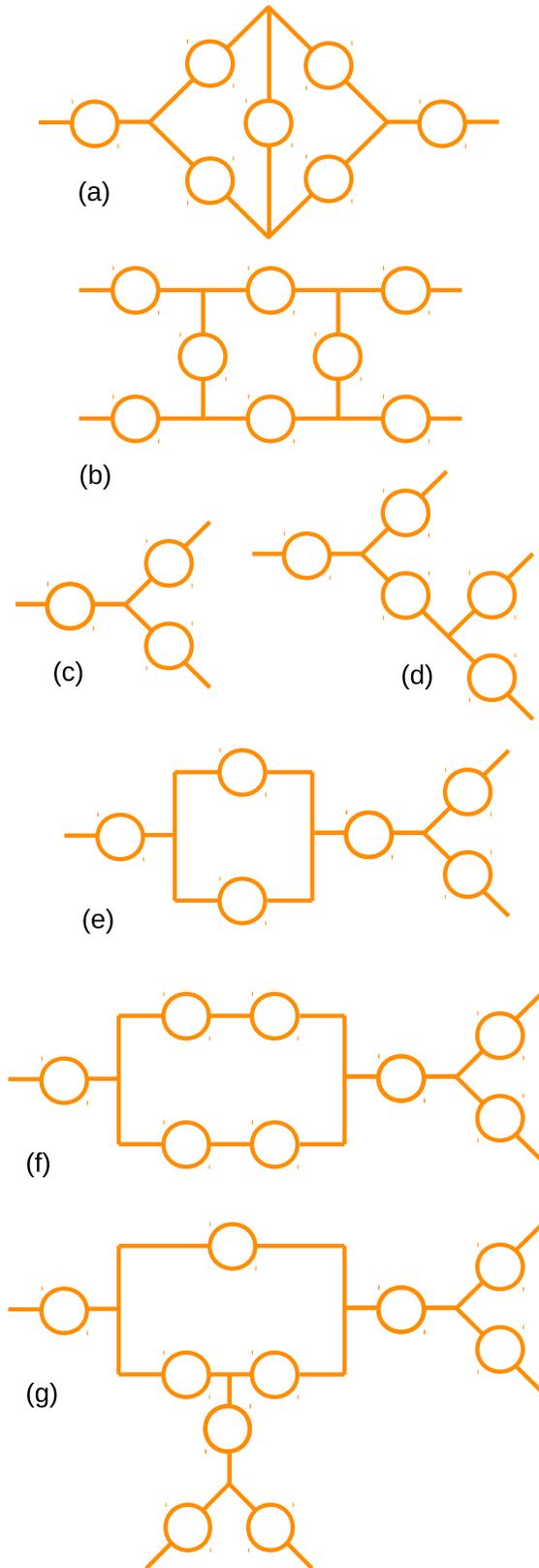


Figure 8. Examples of network topologies: (a-d) are irreducible for parseq algorithm, (c,d) are reducible for tree algorithm, (e-g) require combination of algorithms parseq+tree.

TABLE I. PARAMETERS OF TEST NETWORKS

network	nodes	edges	compressors	psets
N1	100	111	4	2
N2	931	1047	7	4
N3	4466	5362	25	6

In Table II, we compare the performance of the implemented algorithms (strategy “stable”) with the performance of the solver with standard settings. For each network in the test set two scenarios are considered, with different numerical values of set points for input pressures and output flows and compressor/regulator *SM*, *SPO* settings. Divergent cases are marked as ‘div’. The number of iterations (iter.) and timing (t) are given. The simulation is performed on a 3 GHz Intel i7 CPU 8 GB RAM workstation.

All scenarios are tested both with free and advanced compressor models. We see that the standard solver provides worse convergence and even diverges in certain scenarios. Some scenarios show divergence already for free compressors, some diverge on advanced ones only. The new algorithm converges in all cases, in agreement with its theoretical properties. We also see that the table lookup implemented for advanced modeling has a negligible computational overhead in strategy “stable”.

TABLE II. COMPARISON OF THE ALGORITHMS

scenario	solver_strategy							
	standard				stable			
	free		advanced		free		advanced	
	iter.	t, sec	iter.	t, sec	iter.	t, sec	iter.	t, sec
N1S1	3	0.01	32	0.12	2	0.01	2	0.01
N1S2	57	0.17	70	0.21	11	0.03	4	0.02
N2S1	11	0.27	19	0.64	12	0.31	12	0.37
N2S2	div	–	div	–	13	0.36	15	0.48
N3S1	div	–	div	–	26	3.3	23	3.5
N3S2	47	6.5	div	–	26	3.3	24	3.6

TABLE III. NODES / EDGES COUNT FOR REDUCTION ALGORITHMS

network	orig	clean	clean+ parseq	clean+ parseq+tree
N1	100 / 111	37 / 38	10 / 9	2 / 1
N2	931 / 1047	504 / 514	289 / 295	15 / 17
N3	4466 / 5362	1755 / 1843	1012 / 1056	46 / 62



Figure 9. Network N1 after topological reduction.

The network N1, whose original structure is displayed on Figure 1, is ultimately reduced to a single element directly connecting two P_{set} nodes, see Figure 9. The element equation immediately allows to find the flow through the element: $Q(P_{set1}, P_{set2})$. All Q_{set} nodes are moved by the algorithm to the output node, into P_{set2} . The value $Q(P_{set1}, P_{set2}) - Q_{set}$ defines the flow sourced / sinked in this node. The inverse reconstruction procedure will redistribute the cumulative Q_{set} to the actual positions.

In Table III, the efficiency of the topological reduction algorithms applied to the test networks is presented. Already the first step (clean) brings significant reduction, due to the abundance of valves and shortcuts in the network. Further steps bring a significant reduction factor ~ 70 . Therefore, the reduced networks will be solved very fast or will have an explicit solution, like N1 example above. However, one should not forget that the reduction procedure includes operations on the data fields, large tables and is computationally intensive.

Estimation shows that topological reduction requires $O(N_{it1}N_{px}^2N_{elem})$ operations and $O(N_{px}^2N_{elem})$ memory, where N_{it1} is an average number of iterations per pixel necessary to combine two neighbor elements, N_{px} is the resolution of element diagrams, $N_{elem} = N + E$ is the total number of elements. The practical values are $N_{it1} \sim 1-10$, $N_{px} \sim 10-100$. The direct solution also has an empirical estimation of $O(N_{it2}b^2N_{elem})$ operations, with $N_{it2} \sim 10-100$, $b \sim 10-100$. Here N_{it2} is the number of iterations of the non-linear solver, bandwidth parameter b is related with the sparsity of the system. The inverse reconstruction requires only lookup operations on stored reduction history and is performed much faster than the forward reduction algorithm.

According to this estimation, the topological reduction is a bit faster than the direct solution, but of the same order of magnitude. Both procedures are linearly scaled with N_{elem} . It is important that the topological reduction of certain subsystems can be performed once if their parameterization will not be further changed. Also, the reduction of the system leads to elimination of degenerate degrees of freedom and increases the stability of solution procedure as a whole.

VIII. CONCLUSION AND FURTHER PLANS

A mathematical modeling of gas compressors with their individually calibrated physical profiles has been presented. The measured profiles are converted to an explicitly resolved form using the algorithms inspired by computer graphics. The control element equation for a free compressor has been extended by a lookup function representing the working region of the advanced compressor.

The resulting equation possesses the desired signature of derivatives necessary for global non-degeneracy of the Jacobi matrix. Therefore, the globally convergent algorithm earlier developed for the solution of network problems is also applicable

for advanced modeling, with a negligible computational overhead. The efficiency of the approach has been demonstrated for a number of real-life network scenarios. The algorithm significantly overperforms a standard Newtonian solver in terms of stability, number of iterations and computational time.

The algorithms for topological reduction of compressor stations with parallel and sequential connections have been developed. The algorithms combine the lookup tables for individual compressors into a cumulative working diagram, representing the behavior of compressor station as a whole. This allows to eliminate the internal degrees of freedom, further improving numerical stability of the solver.

The topological reduction has been extended by a tree reduction algorithm and its applicability to generic networks of arbitrary elements has been investigated. Estimations of the efficiency of the reduction procedure on realistic network problems shows a potential of significant simplification, with a reduction factor about 70. In certain cases the network appears to be reducible to a single element, thus the solution by the non-linear solver in these cases is actually not needed and the result can be completely reconstructed by simple algorithms from the reduction history.

Our further plans include the implementation of the algorithms with GPU parallelization, extension of compressor profiles with dependencies on temperature and gas composition and a special consideration for nearly singular element equations.

ACKNOWLEDGMENT

We are grateful to the organizers and participants of INFOCOMP 2017 conference for fruitful discussions. This work is supported by German Federal Ministry for Economic Affairs and Energy, project BMWI-0324019A, MathEnergy: Mathematical Key Technologies for Evolving Energy Grids. This material is also based upon work supported by the German Bundesland North Rhine-Westphalia using fundings from the European Regional Development Fund, grant Nr. EFRE-0800063, project ES-FLEX-INFRA.

REFERENCES

- [1] T. Clees, I. Nikitin, and L. Nikitina, "Advanced Modeling of Gas Compressors for Globally Convergent Stationary Network Solvers", in Proc. INFOCOMP 2017, The Seventh International Conference on Advanced Communications and Computation, June 25-29, 2017, Venice, Italy, IARIA, 2017, pp. 52-57.
- [2] T. Clees, I. Nikitin, and L. Nikitina, "Making Network Solvers Globally Convergent", Advances in Intelligent Systems and Computing, vol. 676, 2017, pp. 140-153.
- [3] T. Clees, N. Hornung, I. Nikitin, and L. Nikitina, "A Globally Convergent Method for Generalized Resistive Systems and its Application to Stationary Problems in Gas Transport Networks", In Proc. SIMULTECH 2016, SCITEPRESS, 2016, pp. 64-70.
- [4] T. Clees et al., "MYNTS: Multi-physics NeTwork Simulator", In Proc. SIMULTECH 2016, SCITEPRESS, 2016, pp. 179-186.
- [5] M. Schmidt, M. C. Steinbach, and B. M. Willert, "High detail stationary optimization models for gas networks", Optimization and Engineering, vol. 16, num.1, 2015, pp. 131-164.
- [6] J. Mischner, H.G. Fasold, and K. Kadner, System-planning basics of gas supply, Oldenbourg Industrieverlag GmbH, 2011 (in German).
- [7] J. Nikuradse, "Laws of flow in rough pipes", NACA Technical Memorandum 1292, Washington, 1950.

- [8] C. F. Colebrook and C. M. White, "Experiments with Fluid Friction in Roughened Pipes", in Proc. of the Royal Society of London, Series A, Mathematical and Physical Sciences, vol. 161, num. 906, 1937, pp. 367-381.
- [9] J. Saleh, ed., Fluid Flow Handbook, McGraw-Hill Handbooks, McGraw-Hill, 2002.
- [10] DIN EN ISO 12213-2: Natural Gas – Calculation of compression factor, European Committee for Standardization, 2010.
- [11] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming", Mathematical Programming, vol. 106, num. 1, 2006, pp. 25-57.
- [12] R. Fletcher, Practical Methods of Optimization, Wiley, 2013.
- [13] M. Schmidt, M. C. Steinbach, and B. M. Willert, "High detail stationary optimization models for gas networks: validation and results", Optimization and Engineering online, 2015, DOI: 10.1007/s11081-015-9300-3
- [14] A. Griewank, J.-U. Bernt, M. Radons, and T. Streubel, "Solving piecewise linear systems in abs-normal form", Linear Algebra and its Applications, vol. 471, 2015, pp. 500-530.
- [15] D. Eppstein, "Parallel recognition of series-parallel graphs", Information and Computation, vol. 98, 1992, pp. 41-55.
- [16] N. M. Korneyenko, "Combinatorial algorithms on a class of graphs", Discrete Applied Mathematics, vol. 54, 1994, pp. 215-217.
- [17] K. Takamizawa, T. Nishizeki, and N. Saito, "Linear-time computability of combinatorial problems on series-parallel graphs", JACM, vol. 29, 1982, pp. 623-641.
- [18] M. W. Bern, E. L. Lawler, and A. L. Wong, "Linear-time Computation of Optimal Subgraphs of Decomposable Graphs", J. Algorithms, vol. 8, 1987, pp. 216-235.
- [19] T. Kikuno, N. Yoshida, and Y. Kakuda, "A Linear Algorithm for the Domination Number of a Series Parallel Graph", Disc. Appl. Math., vol. 5, 1983, pp. 299-311.
- [20] M. Hazewinkel, ed., "Green function", Encyclopedia of Mathematics, Springer 2001.

Scientific Challenges in Archaeology

Is modern computer science ready for archaeology?

Lutz Schubert
IOMI, University of Ulm
Ulm, Germany
lutz.schubert@uni-ulm.de

Gill Hunt
Hunt Lancaster
Lancaster, UK
ghunt@huntlancaster.co.uk

Keith Jeffery
Keith G Jeffery Consultants
Faringdon, UK
keith.jeffery@keithjefferyconsultants.co.uk

Abstract—Digital Humanities is receiving growing interest in academia. However, in most cases it is understood as teaching computer science in humanities, much rather than actually merging the subjects. In fact, most computer scientists regard humanities as a “trivial” challenge without “hard, scientific” problems. However, many areas in humanities pose hard challenges to which current computer science cannot provide sufficient solution approaches. In this paper, we examine the specific scientific problems in archaeology that pose new hard and interesting challenges to computer science.

Keywords – advanced applications; archaeology; high performance computing; physics; simulation; network analysis; social networks; agent systems; theoretical computer science.

I. INTRODUCTION

In this paper we examine the challenges presented at the INFOCOMP conference in 2017 [1] and expand them with new insights and a concrete case study:

Dealing with one of the oldest subjects, archaeology is one of the youngest scientific fields, which developed really only in the late 19th century [2]. To many people it comes as a surprise how methodological and scientific such a discipline is that seemingly still works with shovels, pen and paper. In fact, archaeology is an early adopter to many new technologies, such as radiocarbon dating, remote sensing, LIDAR, photogrammetry etc. [3][4].

Yet what computational tasks may arise from a field that deals with interpretation much more than with hard facts? Why should statistical methods, databases and current analytical tools be insufficient? Let alone pose new problems to computer science?

To illustrate the problems faced by archaeology and the types of analysis typically performed, it is best to examine a typical, yet slightly more baffling case in archaeology (Section II). We will use this concrete case study to elaborate the analytical tasks of archaeology and examine the scientific foundations it builds up upon. We will analyse each of these methods in light of current computer science and its deficiencies, as well as potential solution approaches. We

cannot propose a solution here: there is none as yet! All these challenges remain to be addressed by computer science and for this, more interdisciplinary research is needed. We will discuss this briefly in Section III.



Figure 1. Archaeologist mapping the layout of a trench. Source: Wikipedia.

II. CASE STUDY: THE MYSTICAL COIN



Figure 2. The excavation site at Okinawa. Source: Uruma City Educational Board.

In 2016, a group of 10 roundish, metal objects was found in Okinawa, Japan, 1m deep in the ground among a set of symmetrically shaped, roughly geometrically aligned set of

stones [5]. To be precise, two types of shapes and geometries could be observed, but we will ignore this for simplicity sake. In Figure 2, you will probably recognise the ground plan of a building right away, and you may suppose that the metal plates are coins – but bear in mind that Japanese coinage in the past was quite distinctive (see Figure 8).

But we are already jumping ahead of ourselves: past? We are only talking about something in the ground! Building? Aligned stones! Why dig here in the first instance?

A. Geophysics

Many readers will have visited one archaeological site or the other and will probably have a mental image such as of the Acropolis (Figure 3), yet this is obviously basing on decades of reconstruction.

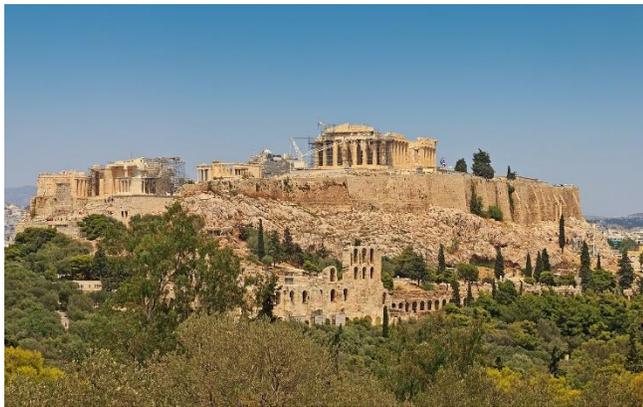


Figure 3. The Acropolis of Athens. Source: Savin, Wikimedia Commons.

Instead, archaeologists are happy if some indicator of the site is visible from the ground (see Figure 4). The most typical approach even nowadays still is to probe the ground at a “promising” location, viz. basing on historical records, sporadic finds or indeed indicators on the ground.



Figure 4. Archaeological site on Samothraki, GR. Try to find the buildings. Photo by the author.

Over recent years, other methods have evolved to detect buildings and structures that were otherwise covered up or undetectable. One such method, LiDAR recently just made the headlines, when the technique helped to discover a Maya

“Megalopolis” in the jungles of Guatemala [6]. LiDAR bases on the capability of laser to penetrate the green foliage, thus displaying the elevation otherwise hidden by the trees.

Another, slightly less known technique exploits the fact that the density, magnetic or electric characteristics are different between materials. Thus, by applying a current, or injecting radar or sound waves into the ground, we can measure different “replies” depending on the material properties and the depth (distance, current, run-time) investigated. GPR (Ground Penetrating Radar) is one such technique, which is also exploited for oil exploration [7].

Such techniques have the major advantage that they are faster than a test dig and do not destroy the evidence in the ground. In other words, the archaeologists can safely explore the area to decide on where and how to excavate. On the negative side, the technique is expensive and still time consuming enough to not be executed on a more global scale, but is more localised.

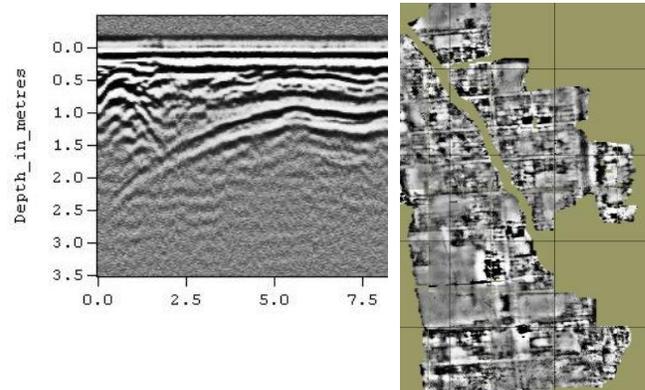


Figure 5. Ground Penetrating Radar (left) and map resulting from a resistance survey (right). Source: Wikipedia.

Scientific Challenge(s). The technique can only observe differences in material properties and allows (so far) little information about the material itself (see Figure 5, left). In general, the runtime evaluation is already computationally intensive [7]. What is more, however, there is still a wide gap to better assess the types of material in the ground and to assess the results in the first instance (see Figure 5, right).

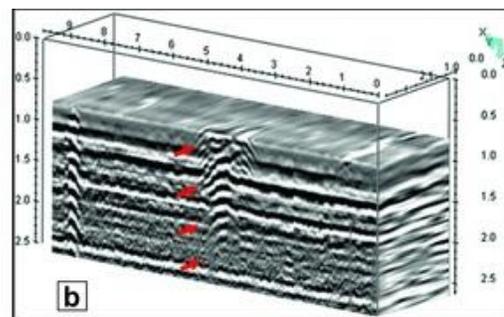


Figure 6. GPR survey of a Roman Stadium floor in Turkey. Source: [8].

Geology can provide additional information about the material constitution in the respective area. This provides a

first insight into the expected conductive / reflective behaviour of the ground material (as is used in oil exploration in fact), and thus any deviation is highly indicative.

At the moment, computation of the results generates a pseudo 3d image that indicates the depth of material changes over the area covered (see Figure 6), but requires a considerable amount of interpretation.

In order to improve interpretation of the data, the different layouts and principle material characteristics need to be put into relationship with regional / context information. The granularity of the technique implies that only structures of larger granularity (typically buildings) can be identified. Nonetheless, as with any human (or in fact animal) artefacts, non-regarding all past and present standardisation efforts, individual structures will differ from each other – be it only for aesthetic preferences. In addition to this, local circumstances, as well as the differences in collapse and deterioration will lead to strong individual deviations.

This deviation needs to be taken into consideration when trying to interpret the structures under ground: changes in material properties may(!) represent structural elements which need to be compared against possible structures in the region – bearing the regional and individual differences in mind. Physical collapse as well as human influence in destruction (both see below) thereby equally play a role – we will come back to that later again.

Matching the 3d information needs a different way of representing and matching 3d structures than currently exists: the shape is not only rotation and scale invariant, but may also change considerably between instances – not only in terms of different collapse and destruction, but also in terms of individual builders' preferences.

Another way with related challenges consists in “space archaeology” [9], i.e., analysis of satellite images for potential features on the ground. Considering the amount of images and the details that could mark a site, this exceeds current image analysis techniques by far: both in terms of identifying the features, as well as the speed of execution.

B. Stratigraphy?

Though analysis of the GPR and satellite data may have given us an indicator for the site to dig, we still need somehow assign a date to the location:

Anyone who has ever dug a hole will have noticed that the constitution of the ground changes in sometimes clearly visible layers (or “strata”). Archaeologists (and in fact all geoscientists) can use this information to measure time – or more correctly, sequences of actions. To illustrate this, think about your own trashcan (see Figure 7): as you fill it over the days, it will accumulate garbage that reflects your past actions: at the lowest layer, the bottom of the trashcan, you can find whatever you threw away the very first day after the trashcan was emptied – let us say it's a set of different organic matter from a nice meal you prepared for friends. The next day you cleaned your house and threw away dirt from the vacuum cleaner, and assembled dirt from around the house – you may not even have eaten a lot, but only had

some leftovers from the previous day, which generated little extra trash. The day after you threw away the old broken cup and some ugly ceramics that was gathering dust in the cupboard... and so on.



Figure 7. Stratigraphic layers in a trashcan. Source: Canterbury Archaeological Trust Ltd.

If you could cut your trashcan in half and look at it from the side, you could now see distinctive layers of different types of trash – each reflecting your daily activities and the garbage you produced. Of course, you will object that these layers are hardly that different over a normal week, which is one of the reasons that only major changes are actually observable (as changes) in the archaeological context. Since these layers decompose and compress, a meter can represent hundreds of years, or as Jared Diamond put it

“we can travel in time, it is just in a different direction than we expected.”

Obviously, the relationship is not as simple as x meters down equals y years: strata build up differently in different regions (think forest floor vs. desert) and human (and animal) interference will change the layout, such as when digging a hole and filling it up with material from another location – we will look at this again later, in the context of “big data” below.

For now, let us assume that this gives us sufficient information to date the set of metal objects to the 15th century. Where was the computer science challenge in this?

Scientific Challenge(s). How material builds up and behaves over time is still mostly unknown. The information available so far bases mostly on experimental data and experience. In fact, chemical properties, composting, even compression of organic material is still too complex for modern computer simulations. Though we can model air and water flow, electro-chemical properties and deformation of metal etc. we still struggle with composite material that combines different properties and, which is worse, reacts with its environment and thereby changes its properties.

As computer scientists, we are used to *precise* computation – in fact, much algorithmic work is invested into increasing the precision without increasing the computation time. However, whilst this is of scientific interest, it is hardly of any practical value. Many computer scientists before have already noted that we could potentially

achieve more if we would change the way we are asking questions – if we invest into “what” we want to achieve, rather than “how”. The EC Cloud Computing Expert Group published one report to the end that software engineering needs to refocus the way we are thinking about problems [10]. Similarly, the whole area of “imprecise computing”, as a rising IT field, is investigating whether we cannot achieve the same results with less precision [11].

What is more, the timescale of a typical simulation ranges in the order of milliseconds (such as for protein folding) to minutes (such as in airflow simulation). The longest time-span is probably in the order days with weather simulations, but bear in mind that this already belongs in the realm of imprecise and statistical computing.

Multi-level computing is still on the rise and problems such as the Virtual Physiological Human [12] cannot be solved before we cannot reduce the computational complexity. Organic composting, compression and composite material, though all highly important for building management, engineering etc. are so far only scientifically investigated by archaeology and build up on human expertise and experience.

C. Multidisciplinary Statistical Big Data

Until now, we identified the site itself and have a rough feeling for the time to which the metal plates belonged. As we will discuss in more details below, timing just by stratigraphy is insufficient as it differs too strongly between regions and circumstances.

One of the most frequent and most challenging tasks in archaeology consists in data analysis (and as we shall see, reasoning): in our example we have identified so far a location, the structure itself and a set of round metal objects. Analysing them goes far beyond the traditional material probing (though this can help dating). Primarily, any interpretation in archaeology is trying to link data with known facts. Let us concentrate on the building itself first:

We have known parameters: location, rough dimensions and ground plan, as well as depth (and therefore indicative age). With this, we can start looking for registries, historical records, architectural analysis etc. Other potential information involve the geological source of the stones, the shape and methods in which the stones were shaped (as observable by rough and microscopic surface analysis), the overall layout (as in cultural building traditions), any signatures, stamps, inscriptions and of course all other objects found on site. Archaeology is one of the most interdisciplinary scientific fields and combines information from a multitude of sources – we will come across this issue repeatedly in this text. For now, however, let us assume that historical records of the respective building exist – in this case, a Japanese castle from the 14th century.

Scientific Challenge(s). Cross-relating such information is a typical task and requires searching through different sources, ranging from geology, over climatology to anthropology, architecture, biology etc. Archaeology thereby has way less hard data at its disposal than other scientific fields – already the constant changes in radiocarbon-dating [13] or the constant reinterpretation of genetic data [14] is

witness to the instability. In most cases, these re-examinations were in fact triggered by archaeologists, when the data generated by the so-called “hard scientific methods” was in conflict with the archaeological evidence or contradictory to reasoning.

The main challenges is however not so much the scope and interdisciplinarity, as one might expect, but the fact that most of the data is not concrete and that there can be only “interpretative” links, as the actual “facts” may change at any time. Just like in other scientific fields, once enough evidence is gathered that is in conflict with current thinking, the current facts need to be rethought. In the context of humanities, however, the actual subject (humans) are way more complex and facts considerably less stable.

In other words, the main challenge is to reason over different data and assess likelihood of certain events or situations. As we shall see, a simple stochastic reasoner is insufficient, as we must also consider stability of a belief and thus its impact on the overall conclusions, should it change.

Data mining hence needs to be extended with 3 aspects: interpretation (as any social data is subject to interpretation rather than a simple hard fact), reasoning (see below) and likelihood. In terms of likelihood, we need to distinguish between direct facts (i.e., the evidence gathered from whichever means – be that the stones or the metal plates on site, or the genome sequence from a human bone) and inferred facts and their relationship. For example, the *best* evidence suggests that the rock assemblage was indeed a Japanese castle (given the location and the rough age), but the documents may be false, the site may have been reused within a short timespan etc. In almost all archaeological cases no historical data exists in the first instance, thus making the interpretation even more subject to likelihood based on inference and assumptions.

The archaeological pictures (and in fact our understanding of ourselves as humans) gets constantly updated, as we gather more knowledge and gain more insight. In every step, all knowledge acquired needs to be considered and revised – potentially affecting *all* data.

Let us look back at the metal plates found in what we know presume to be a Japanese castle of the 14th century. By logical inference, we may first assume that the metal plates must be as old as the castle, but we have to be careful about this interpretation: the castle may have been built in the 14th century, but it was in use until the 15th century. Even then, it only slowly disintegrated and thus the coverage of dirt may indeed be from the 18th century and the metal plates of according age. In other words, the finds can only be as old, as the first layer on top of it – at the same time, it cannot be older than the first layer underneath it. Bear in mind that this applies to the age of last use of the object, not to the age of the object itself.

In this specific case, the evidence suggests indeed that the objects were or discarded some time in the 14th or 15th century. But what does that make them?

D. Advanced Image Recognition

As already noted in “Geophysics” above, one of the typical challenges consists in interpreting evidence –

visually. Though more elaborate analysis methods exist, they typically base on material constitution, chemical consistency etc. but only visual analysis helps for interpreting shapes and shapes can be interpreted through similarity: in rare cases are two objects identical – in particular if manufactured by hand. Yet we can infer producer (artist) and cultural traits from artistic decoration and shape [15] – for example, the “art nouveau” period left a distinctive mark in visual cues and it is pretty easy to distinguish individual artistic schools. Such classification is subject to a lot of research in image recognition, yet in our case here, we not only have to recognize styles, we also have to deal with deterioration and destruction.

In the case of our set of metal plates, the first association by inference was that these may have served as armor plates, as local coinage would have had a distinctive square hole (see Figure 8).



Figure 8. Examples of Japanese coinage from the 14th-17th century. Source: Wikipedia.

It was sheer luck that one of the archaeologists recognized a similarity with Roman coins based on experience gained in Egypt. What makes the challenge much harder is the fact that not only knowledge about the other context (here: Roman) is needed, but also that artefacts deteriorate. In case of the coin, Figure 9 depicts how the coin was found in Japan (left) and how it could have looked like if it would have been preserved better (right).



Figure 9. The Roman coin found on the Japanese site (left, Source: Uruma City Educational Board) and a well preserved version of (probably) the same coin (right, Source: www.wildwinds.com).

Scientific Challenge(s). As we have already noted multiple times, identification of finds, their classification and timing requires a big data comparison over multiple different sources. The case here clearly illustrates how ignorance of all information can lead to misinterpretation of the artefacts and thus to misinterpretation of whole region and its past. We will see later how the (correct) interpretation is actually not very helpful in this case.

Years are spent in analysing the finds from a site and even with the best archaeologists, mistakes in interpretation can happen. A computer aided interpretation process could greatly support this endeavour but would require that image recognition is expanded to consider deterioration processes. Much like we will discuss below in “Physics” and in “Geophysics” above, these processes can be modelled to a certain degree.

Differences in the manufacturing process (such as stamping coins) can already to a certain degree be analysed, yet the process is still very crude and does not really allow to compare coins of different origin. Analysis of microabrasions, experimental archaeology and simple experience so far help much more than the analysis capabilities of modern software. Again, combination of approaches are needed to expand the algorithmic quality.

E. 3d Recognition and Matching

More typically, though not the case in our example, archaeological finds are in fragments, such as broken pieces of ceramics, or the famous Lionman, which consists of more than 300 fragments of ivory and still is not complete – the smallest of the fragments is thereby only a few millimetres.

Next to the general layout of the finds, the actual material and shape of the fragments themselves provide indicators for their relationship. Consider the various forms of pottery that can be found in archaeology: shape, material and texture, respectively decoration are good indicators as to whether two sherds may have belonged to the same object. This also applies to (human) bones, larger sculptures etc.

Generally, parts are missing, scattered, or even archived in a completely different city / country due to different excavation processes, movement after excavation etc. Furthermore, due to the vast amount of similar fragments, identification of corresponding parts is close to impossible.

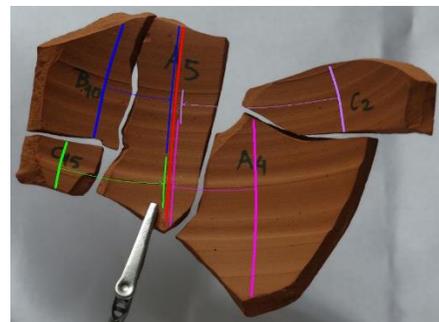


Figure 10. Matching sherds on basis of profile information [16].

Scientific Challenge(s). In itself, this is a considerable big data process where multiple factors need to be compared and cross-correlated to identify potential matches. These in turn will have to be matched in shape and against types of objects. Ideally, the fragments touch and thus have a common breakage area. Though this may sound like “just” a fitting task, one needs to consider (a) the amount and size of fits and (b) that natural processes change the breakage area by smoothing and reducing it etc., so that no perfect fit can

be achieved. Such processes have to be taken into account when asserting whether two fragments match [17].

In most cases, no 3d models are available, let alone sufficient details to attempt a match in the first instance. As the number of available models increases, so does the complexity to match all the available finds – but even just within a single excavation, the effort is considerable.

As noted, though, in our example no such matching needs to take place.

F. Physics & Material Deterioration

We already mentioned in the context of Geophysics, above, that physical (there: biochemical) processes need to be understood when examining a site or the artefacts found therein. Of specific interest here are thereby the deterioration processes from exposure to the elements – this equally includes rusting of the metals objects, stratigraphic built-up, as well as structural collapse.

In the case of our castle, it is not only covered up by increasing layers of compost, loose dirt and rubble. The building probably also collapsed due to structural weakness and, what is more, human intervention: people always have and always will reuse buildings for new structures, as a material source, or just use the space for other purposes, thereby flattening the remainders. This poses extremely hard challenges on reconstructing the processes and simple rigid body physics simulations are insufficient to take all these factors into account.

It is however not only structural analysis that can benefit from physics simulation: humans and animals having been killed violently and / or moved after death will end up in certain positions and orientation. For example, skeletons in the Tollense valley have been moved by water slides and thus ended up in a collective heap [18]. Knowing the shape of the land, the flow of water and intensity of rainfall allows reconstructing where the bodies originated from, and (to a degree) their original positions. As the process is irreversible, this is not entirely possible – but the order of skeletons already indicates how they must have been flooded down the hill. Notably, the state of decomposition makes a major difference with this respect.

Related to this, marks in the skeleton give an indicator for strength and direction of a blow or projectile. Arrow-heads embedded in bones tell something about the position of the opponents relative to each other, but also about how the weapon was used and the force that the respective weapon can transmit. Human factors have to be taken into consideration, such as whether the force could be created by muscle strength (spear) or whether additional means would have been needed (bow). Given e.g., the Tollense layout, a reconstruction of the event can be attempted.

Scientific Challenge(s). As noted, rigid body physics simulations are good to get some feeling for how buildings can collapse, and thus good for educative purposes, but they are far from providing insight into the processes that actually took place on site. Just as with geophysics, what is needed is multi-level simulations in order to take all the different factors into account.

By nature, however, physics is a chaotic process with little influences leading to major changes in the outcome. Therefore, it is highly unlikely that we can recreate all the circumstances that led to a specific constellation as found in an archaeological site (or in fact any site). Archaeologists essentially invert physics and try to find logical explanations for which forces may have acted to result in this outcome.

Physics inversion is mathematically impossible, but stochastic relationships between forces and outcomes are possible [19][20][21]. The shape and layout of the elements in the heap allow reasoning over the possible original structures. By comparing these possibilities with existing, similar ones, we can make even reasonable assumptions about the factors that led to the final distribution.

Such methods essentially can only be used for verifying (respectively falsifying) certain assumptions, such as that the likelihood of a specific distribution of stones can be the result of natural phenomena, or that the distribution of food remainders has not been tampered with [19][22].

Notably, we need an assessment for the forces that may have acted on the site – one of the most difficult to measure and analyse being human intervention.

With all the information gathered so far we know now that the metal plate is indeed Roman, and given the deterioration, as well as the expected initial shape, we can assume that the coin is from the Roman Emperor Constantine. We also know that the coin must have been lost some time in the 14th or 15th century. Now let us apply our big data analysis again: Emperor Constantine lived from 272 to 337 AD – certainly way before our Japanese castle was built. His coins may have been in use past his reign, but since every emperor decreed new coins, it is unlikely that this lasted for very long. Obviously, the coins did not get in full disuse, yet at the 14th century, Roman coinage was not a currency anymore. Though the Roman empire still existed at this point, it had lost most of its influence and was quickly succumbing to the Arabian conquest (see Figure 11).



Figure 11. The Roman empire around the 14th century. Source: Wikipedia.

Which leads us to a simple question: how did a more than 1.000 year old coin end up in a place more than 10.000 km away from its origin (see Figure 12)? Japan was never part of the Roman empire and had its own coinage, but some human

contact must obviously have taken place: we have hard scientific evidence in the undeniable coin's existence and location.

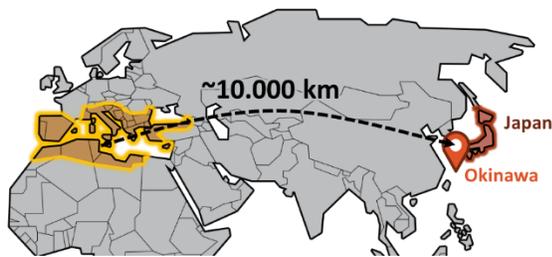


Figure 12. Map of Eurasia showing the main Roman empire and Japan.

What can we say about the human agents that must have been involved?

G. Simulating Human Behaviour

Archaeology is about humans: how they lived, what they have done, when and why. However, in an illiterate society, ways of thinking leave no traces and even in literate societies, written evidence should not be confused with facts [3]. The challenge for archaeology therefore consists in relating finds to potential behaviour, intentions and way of thinking. Some of this behaviour is obvious and straight-forward: a ceramic pot indicates that (a) someone was there to leave the pot behind and (b) someone made the pot. However, was the pot used as a domestic item, was it an item of worship, was it just decorative, was it discarded right away? All this cannot be gathered from the pot alone.

As seen (data mining, above), a considerable amount of information has to be cross-linked. What is more, though, is that human behaviour, intentions and beliefs, capabilities and knowledge etc. stand at the middle of the explanation chain and form the basis for any conjecture. As indicated above, this can obviously take different levels:

Presence. Straight-forward, remains are “just” indicators for human presence and actions, such as that someone must have brought the find to the location, must have made it etc. Notably, not always is a find clearly of human origin, as e.g., is the case with some Palaeolithic “tools” [23]. This is the level of direct archaeological evidence.

Capabilities. At an intermediate level of complexity, human capabilities must be taken into consideration. This defines whether it was e.g., possible to reach a location, build a structure etc. How humans reached the American continent is one such unsolved question. At this level we talk about the assumptions that can be substantiated by archaeological evidence (existence of boats), but not fully proven.

Belief and Intention. At the most complex level we need to argue over belief and actions that are behind the evidence. It is a frequent cliché that archaeologists classify any evidence without clear functionality as “ritualistic”. Indeed, it is difficult to assess the intention of an object that has no comparison in modern context. At this level, all “evidence” is pure conjecture and may change on basis of new theories.

Whereas knowledge at level 1 and partially at level 2 falls clearly into big data management, i.e., cross-checking

facts, most of level 2 and in particular level 3 are conjecture and base on logical possibilities alone. Aspects such as movement of peoples require that the behaviour is simulated and the likelihood assessed on basis of this simulation.

Scientific Challenge(s). Even the best swarm simulation software cannot accurately model human behaviour beyond simple crowd movement. The typical approach consists in agent based simulations, which model multiple entities and their interactions on a simplified level [24]. There is a considerable amount of criticism of these models, as they must be incomplete and error prone – it is currently not possible to appropriately simulate how even just a small settlement would behave [25].

Human behaviour is complex and cannot be easily abstracted, so a major question relates to which human aspects have to be modelled in the first instance and how. Much can be learned from social network interactions, but care must be taken when applying modern contexts to ancient circumstances, as behaviour and mindset are in constant flux [3]. As indicated in the context of simulating physics, we can however argue about the likelihood of a specific behaviour (see below).

Statistical analyses can reduce the computational effort, even though they have a high error margin. They can help to eliminate *unlikely* situations, such as for the Roman “tourist” in Japan, which would necessitate the according means of travel, communication etc. In [26] the authors suggest an analysis based on throwing angles and strengths to assess the layout of shell middens. This is a highly simplified human behaviour model but already allows for some degree of feasibility assessment.

H. Network Analysis

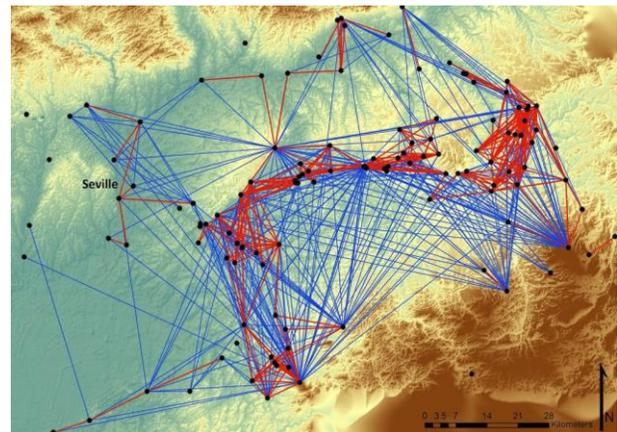


Figure 13. A social network diagram for Facebook. Source: Archaeological Networks.

A special and indirect form of human behaviour analysis consists in social network analysis (SNA). SNA gets a growing interest in archaeological analysis, as it is taken as an objective evaluation of the hard evidence. It belongs into the domain of big data analysis, but given its prominence and growing interest, it deserves its own subsection.

A network analysis assesses the statistical likelihood of two entities being related or connected. It is frequently used

in social analysis to identify social relationships between people all over the world. Typical parameters are thereby simple knowledge of each other such as through Facebook (see Figure 13). Such a network can be analysed in different ways: if only social relationship is observed, clusters of high connectivity depict the relational intensity, i.e., “friends with common friends” – outliers or loose connections between clusters indicate isolated groups with little knowledge of each other. Such a relationship is interesting e.g., for target market analysis, to see which people may most influence each other’s taste.

Frequently, other parameters are used for analysis, such as indeed shopping behaviour: people with similar taste are more likely to buy the same items. In other words: the more your shopping behaviour links to other people, the more likely their taste is comparable to yours. Whilst this works on average, it certainly fails in the specifics, as everyone looking at “personalised recommendations” knows.

In archaeology, other “similarity factors” are typically used, such as the similarity between art forms, genetic relatedness, or in our case, appearance of Roman coins. To be meaningful, the information typically needs to be projected onto a map, ideally also across time. Not surprisingly, we would expect that the occurrence of Roman coins diminishes over distance and time – and in fact, no other occurrence of Roman coins in Japan is known.

Scientific Challenges. SNA is far from a hard analysis tool and there are many problems with such an analysis. Unfortunately, results from SNA are frequently mistaken for facts, leading to potential misinterpretations.

Let us look at some of the base issues with network analysis – specifically in the context of archaeology:

1. Encoding “soft” parameters: similarity and thus potential relationship based on e.g., similarities of shape need to be encoded in a parametric form. Try a network analysis of all different forms of art nouveau and its appearance over the globe and you will notice that defining similarity is a hard task. By using too many parameters, relationship will diminish to nothing and by using too few (e.g., “coin”) you will end up with too many relationships. Depending on who defines the parameters and how, completely different graphs can be produced – and hence completely different interpretations.

2. “Absence of evidence is not evidence of absence”: just because no Roman coins have been recorded in Japan, this does not mean that they may not have been misinterpreted (or simply not found yet). Obviously, the network will immediately change, when more evidence is uncovered, but the reason for evidence being available is manifold and ranges from archaeological accuracy, destruction, investment to preservation conditions.

3. Different analysis methods lead to different results, but their interpretation is once again up to the human user: the obvious connection between the Roman empire and a Japanese castle does not imply that a Roman time traveller lived in a Japanese castle – we will need to turn back to that in “Reasoning” below.

Network analysis in the context of archaeology needs to be extended with the capability to assess the “stability” of the

relationships. In other words: which impact would other analysis methods, new evidence, or change in parameters have on the network shape and in particular on its clustering.

Another, growing trend, is to assign weights to the relationships in the network graph – mostly in order to reflect the transportation or communication cost implied by distance (i.e., when project the graph onto a map). This provides an additional interpretation factor, which relates to the likelihood aspects mentioned above: the more difficult (costly) the connection, the less the impact and hence the weaker the relationship.

In context of the Roman coin, the cost for travelling 10.000 kms at Roman time would be tremendous. However, we thereby neglect a few factors, which we shall turn to next:

I. Reasoning

With the data we have from the Japanese site so far we may infer multiple interpretations, such as that a Roman travelled to Japan to live there and leave the coin as a heritage to his children; that a 2nd century Japanese travelled to Rome and returned with a coin; that a 14th century Japanese travelled through time and space and so on. Though some of these explanations will immediately seem ridiculous, we must note that none of them contradicts the data or interpretation we have so far.

The key point here being that some of the interpretations stand to *reason* that they cannot be correct – such as for the time traveller. In fact, if we consider the data carefully enough, we will even notice that the Roman travelling 10.000 kilometers is highly unlikely given the situation and travelling options at the time. This form of restricting the options (and thus weighing the relationship) is a form of big data analysis, but implies logical constraints and derivations – for example, by applying the average travelling speed at given times as a factor.

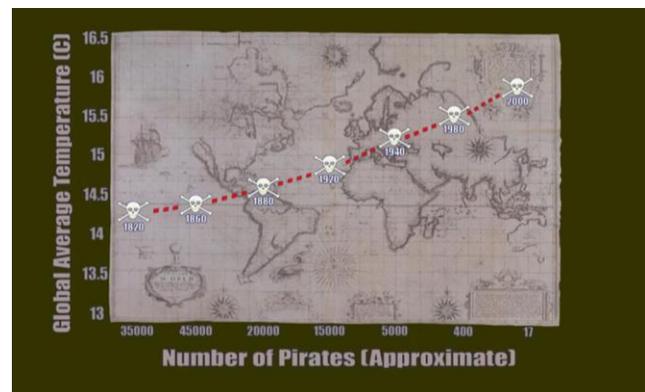


Figure 14. Relationship between numbers of pirates and global average temperature. Source: Forbes.

Modern analysis methods are bad at such behaviour and generally just generate statistical relationships. A classic example of such a “maladjusted” interpretation is the relationship between pirates and global warming: because a correlation between numbers of pirates and global temperature can be observed (see Figure 14) [28]. Since there is no logical check involved, this assumption is fair –

we may even be inclined to take the inversion of the false cause-and-effect chain and thus presume that we need more pirates to counteract global warming. It is only because it is so obviously wrong, that this assumption is not taken up – what, however, if the logical mistake is not so obvious?

Scientific Challenge(s). Reasoning in AI methods is still a new field, but an increasing amount of voices raise concerns about potential misinterpretations of data. Note though that reasoning itself is a type of big data challenge as even though the logical operators are constraint, the number of potential logical assertions is infinite.

Such attempts relate strongly to the domain of semantic reasoning and are still under development, but should not be left aside when interpreting data.

What does all this leave us for interpreting the data at the Japanese site? Obviously, we cannot get a definite answer, but several explanations reach a certain plausibility: (1) there are more Roman coins in Japan that just have not been discovered, (2) the most likely candidate for a transfer route is the silk road, which equally existed in Roman and medieval times. We can only make assumptions about why the coin ended up here and obviously, anything from a collector, over heirloom to an exotic gift may be valid.

Obviously, we can come up with many more alternative routes, but we hopefully managed to reproduce the archaeologist's way of thinking in an algorithmic fashion, as well as showing which challenges this still poses for computer science.

For the sake of completeness, we can identify many more such challenges, such as:

J. Simulating Climate

Climate is constantly changing – not only due to human interference, but also due to the earth's rotation and movement, leading to glacial and hot periods. The implication of such weather changes is obvious and can already be observed today: different plants grow in different climatic zones, animals (and certainly humans) move to different areas, clothing changes etc. In times before Air Conditioning, this hit doubly strong and will have caused (and prevented) massive movement and settlement patterns, following game or reacting to environmental pressure.

Climate completely changes the face of the earth, from rising (and sinking) sea levels to landscapes covered in ice sheets or turned into steppes. These changes leave their marks and are sometimes directly measurable, such as in tree growth (dendrochronology) or remains of marine life in the dessert, respectively vice versa [3] [4].

In the archaeological context climate is only of interest insofar as it influences humans [27]. As such, it is only a contributing factor to Simulating Human Behaviour (see Section II.G) and can serve equally as an explanation, as well as an obstacle. For example, the movement of Homo Sapiens to the American continent is frequently explained by the possibility of a connection between North America and Siberia (the Bering land bridge) [29]. This land bridge could have existed due to a massive amount of water being locked in ice, thus causing the sea-level to sink considerably.

Similarly, the movement of hominins into central Europe from Africa may have been made possible by fluctuations (inter-pluvial arid periods) in the temperature of the Sahara [30].

Climate conditions apparently play a role in any discussion about behaviour influenced by weather, such as clothing, foodstuff etc. Therefore, modelling the weather and in particular the climatic changes over history is a relevant aspect of the argumentation chain related to Simulating Human Behaviour (see Section II.G).

Approaches. It is well-known that weather simulation belongs to the most difficult tasks in advanced applications [31]. While meteorological simulations try to accurately predict local, minute changes in the weather, climate models can be more coarse-grained, identifying patterns of general weather trends over longer periods of time. However, already the overall climatic changes in the glacial and interglacial periods are difficult to predict and not all factors are known. Such models base more on observed factors, such as glacial movements and encapsulated CO₂, than on calculations [4].

Nonetheless, different models are under development [32] and particularly try to provide more local and fine-grained climatic conditions, so as to assess the size and distribution of ice sheets, but also just to predict shorelines, climatic zones etc. Such models can be validated partially against archaeobotanical finds, i.e., seeds that have been preserved under anaerobic conditions.

K. 3d images

3d scanning is a growing field of interest in general, but also more and more archaeologists make use of photogrammetry to document the excavation [33]. There is a high risk that this is considered sufficient documentation, though it cannot replace profile drawings or good maps, but we shall not follow this discussion in this paper.

Generating 3d models from pictures taken in the open field is still time consuming and error prone, where missing pictures can only be identified after generation of the point cloud, which can take days in itself. Since the excavation will have progressed by then, this can lead to considerable problems. Better methods are needed to assess quality and potential gaps right at the time of taking the pictures, and the process in general needs to become more flexible – both require new algorithmic approaches that are highly related to performance optimisation in general.

One should also not ignore the fact that 3d scanning generates massive amount of data (i.e., the 3d points) that so far cannot be easily processed. Identifying an object in 3d space, i.e., which points belong to each other to form an artefact of its own, is still basically impossible. Similar challenges exist in 2d image analysis, where major progress has been made. So far most approaches simply generate a mesh of the whole scan, thus not allowing to (re)move individual objects, let alone perform an analysis on this level.

Since the advent of LIDAR scanning [34], processing of 3d images becomes an important factor for detecting hidden and obscured structures, very similar to identifying hidden structures in geophysical data.

Approaches. So far, most approaches rely on methods from 2d image processing, such as similarity of colour, identification of key features and of their relationship etc., but application in 3d is still very limited – not alone because the size of data is considerably larger (at least from n^2 to n^3).

Google and Microsoft already try to incorporate scans and 3d data from multiple (social) sources, but the sheer amount and computational complexity is still an unsolved challenge. Ideally, however, multiple sources are integrated in scanning, but notably, these will all have to be calibrated individually and the data then has to be cross-correlated first.

Some attempts also try to make use of additional data, such as arising from the accelerometer to pre-assess the quality and usability of the images, but there is no general good solution as yet and the amount of data will only increase.

III. CONCLUSIONS

The list of issues presented in this paper is far from exhaustive but already demonstrates the shortcoming of current computer science methodologies with respect to the needs of archaeology. Specifically, by addressing these challenges and incorporating knowledge from archaeology, the following improvements could be achieved:

- improved geological modelling: archaeology has knowledge about more short-term processes, such as soil deposition and collapse that can be exploited for engineering, city planning etc.;
- better human and agent models: anthropology and archaeology have information about human movement that is not reflected in simulation, thus leading to unrealistic movement and agency models;
- prospecting can benefit from prediction models and material knowledge gained from excavations;
- data mining and big data do not address complexities raised by such interdisciplinary fields as archaeology, which develops such methods for 100 years now;
- statistical analysis is an important field in archaeology and needs to be applied differently for network analysis, clustering etc. The feedback is rarely incorporated (see e.g., [35]);
- structure from motion is constantly being improved through landscape archaeology and field surveys [33] – new more robust methods and better object recognition are still being researched;
- most simulations model time forward from a given situation– in archaeology, time needs to be modelled backwards, i.e., leading from effect to cause, which in turn improves simulation performance and analysis capabilities [26];
- dealing with incomplete data by adding assumption models: archaeology is using methods for this on a daily basis, yet big data still struggles with it;
- both fields need better methods to capture the probability and likelihood of complex data to be correct and to identify logical and improbable errors;

- reasoning needs to improve beyond stochastic data mapping and in particular needs to include the probability that two actions are related. Artificial Intelligence concepts from the 90ies already approach such issues on a limited scale.

Not only can computer science improve archaeology further, but also knowledge from archaeology can help advance computer science capabilities in particular for application in any human-centric simulation or modelling.

We hope that this paper has shown that there remain many challenging tasks for IT in archaeology and that computer science still has many things to learn from the approaches in archaeology. The authors directly contribute to such a collaboration via a dedicated working group in the CAA International (Computer Applications and Quantitative Methods in Archaeology).

REFERENCES

- [1] L. Schubert, K. Jeffery, and G. Hunt, "Challenges for Advanced Applications in Archaeology," In Proceedings of The Seventh International Conference on Advanced Communications and Computation, June 25-29, 2017, Venice, Italy, C.-P. Rückemann, I. Flood, I. Schwerdtfeger, C. Simmendinger, G. Beckett, (Eds.), International Academy, Research, and Industry Association, 2017, pp. 29-34, ISSN: 2308-3484, ISBN: 978-1-61208-567-8, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2017_3_30_60034 [retrieved: March 2018].
- [2] K. Greene and T. Moore, "Archaeology: An Introduction," Routledge, 2010. ISBN: 978-0415496391
- [3] M. Sutton, "Archaeology: The Science of the Human Past," Routledge, July 2015. ISBN: 978-0205572373
- [4] M.J. Aitken, "Science-Based Dating in Archaeology," Routledge, February 2014. ISBN: 978-0582493094
- [5] E. Jozuka, "Ancient Roman coins found in ruined Japanese castle," CNN, 2017 [Online]. Available from: <http://edition.cnn.com/2016/09/27/luxury/ancient-roman-coins-japan/> [retrieved: May 2018]
- [6] T. Clynes, "Ground Penetrating Radar Theory and Applications," National Geographic News, 2018 [Online]. Available at: <https://news.nationalgeographic.com/2018/02/maya-laser-lidar-guatemala-pacunan/> [retrieved: May 2018]
- [7] H.M. Jol (Ed.), "Ground Penetrating Radar Theory and Applications," Elsevier, 2008. ISBN: 978-0-444-53348-7
- [8] V. Karabacak, Ö. Yönlü, E. Dökü, N. Kiyak, E. Altunel, S. Özüdođru, C.C. Yalciner, and H.S. Akyuz, "Analyses of Seismic Deformation at the Kibyra Roman Stadium, Southwest Turkey," *Geoarchaeology*, 28, 2013, pp. 531-543. <https://doi.org/10.1002/geo.21456>
- [9] R. Lasaponara and N. Masini, "Satellite Remote Sensing: A New Tool for Archaeology," Springer, 2012. ISBN: 978-9400796157
- [10] L. Schubert and K. Jeffery, "New Software Engineering Requirements in Clouds and Large-Scale Systems," in *IEEE Cloud Computing*, vol. 2, no. 1, pp. 48-58, Jan.-Feb. 2015. <https://doi.org/10.1109/MCC.2015.13>
- [11] J. M. Cadenas and M. C. Garrido, "Intelligent Data Analysis, Soft Computing and Imperfect Data," in *Soft Computing Based Optimization and Decision Models*, D. A. Pelta and C. Cruz Corona, Eds. Cham: Springer International Publishing, 2018, pp. 25–43. ISBN: 978-3319642857
- [12] M. Viceconti, "The Virtual Physiological Human: Computer Simulation for Integrative Biomedicine: Papers of a Theme

- Issue,” *Philosophical transactions of the Royal Society of London / A: Philosophical transactions of the Royal Society of London*, 2010. <https://doi.org/10.1098/rsta.2010.0096>
- [13] Y. Asscher, S. Weiner, and E. Boaretto, “A new method for extracting the insoluble occluded carbon in archaeological and modern phytoliths: Detection of ^{14}C depleted carbon fraction and implications for radiocarbon dating,” *Journal of Archaeological Science* 78, 2017, pp. 57-65. <https://doi.org/10.1016/j.jas.2016.11.005>
- [14] R. Eisenhofer, A. Anderson, K. Dobney, A. Cooper, and L.S. Weyrich, “Ancient Microbial DNA in Dental Calculus: A New method for Studying Rapid Human Migration Events,” *The Journal of Island and Coastal Archaeology*, 2017, pp. 1-14. <https://doi.org/10.1080/15564894.2017.1382620>
- [15] A. M. W. Hunt, “The Oxford Handbook of Archaeological Ceramic Analysis,” Oxford University Press, 2017. ISBN: 978-0199681532
- [16] M.I. Stamatopoulos and C.-N. Anagnostopoulos, “3D digital reassembling of archaeological ceramic pottery fragments based on their thickness profile,” *ResearchGate* 2016. [Online] Available at: https://www.researchgate.net/publication/301873740_3D_digital_reassembling_of_archaeological_ceramic_pottery_fragments_based_on_their_thickness_profile [retrieved: May 2018]
- [17] F. Stanco, S. Battiato, and G. Gallo, “Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks,” CRC Press, 2011. ISBN: 978-1439821732
- [18] A. Curry, “Slaughter at the bridge: Uncovering a colossal Bronze Age battle,” *Science Magazine* 2016. [Online] Available at: <http://www.sciencemag.org/news/2016/03/slaughter-bridge-uncovering-colossal-bronze-age-battle> [retrieved: May 2018]
- [19] L. Schubert and K. Jeffery, “Interpreting Shell Middens through Physics Simulation,” In *Proceedings of the Computer Applications & Quantitative Methods in Archaeology (CAA) International Conference*, 2017. In print.
- [20] L. Schubert, S. Metz and K. Jeffey, “Inverting Physics – a future for the past,” In *Proceedings of the ICNAAM Conference*, 2017. In print.
- [21] L. Schubert, R. Warden, and T. Guderjan, “Site Reconstruction from Partial Information,” In *Proceedings of the Computer Applications & Quantitative Methods in Archaeology (CAA) International Conference*, 2018. In print.
- [22] L. Schubert, A. Predoi, and K. Jeffery, “Analysing the archaeological context: Reconstructing stratigraphic layers,” In *AIP Conference Proceedings*, 2017. <https://doi.org/10.1063/1.4992196>
- [23] K. Wong, “Wild Monkeys’ Stone ‘Tools’ Force a Rethink of Human Uniqueness,” *Scientific American*, 2016 [Online]. Available at: <https://www.scientificamerican.com/article/wild-monkeys-stone-tools-force-a-rethink-of-human-uniqueness/> [retrieved: May 2018]
- [24] G. Wurzer, K. Kowarik, and H. Reschreiter (eds.), “Agent-based Modeling and Simulation in Archaeology,” Springer 2015. ISBN: 978-3319342825
- [25] D. Thalmann and S.R. Musse, “Crowd Simulation,” Springer 2012. ISBN: 978-1447144496
- [26] L. Schubert and K. Jeffery, “Modelling Physics in Shell Middens,” *Proceedings of the CAA 2017*, in print
- [27] M. Tallavaara, M. Luoto, N. Korhonen, H. Järvinen, and H. Seppä, “Human population dynamics in Europe over the Last Glacial Maximum,” *PNAS*, vol. 112, no. 27, pp. 8232-8237, July 2015. <https://doi.org/10.1073/pnas.1503784112>
- [28] E. Andersen, “True Fact: The Lack of Pirates Is Causing Global Warming,” *Forbes Online*, 2012. Available at: <https://www.forbes.com/sites/erikaandersen/2012/03/23/true-fact-the-lack-of-pirates-is-causing-global-warming/> [Retrieved: May 2018]
- [29] S. Armstrong Elias, “First Americans Lived on Bering Land Bridge for Thousands of Years,” *Scientific American* 2016 [Online]. Available at: <https://www.scientificamerican.com/article/first-americans-lived-on-bering-land-bridge-for-thousands-of-years/> [retrieved: May 2018]
- [30] W.J. Burroughs, “Climate Change in Prehistory: the end of the reign of chaos,” Cambridge University Press, 2007. ISBN: 978-0521070102
- [31] U. Gähde, S. Hartmann, and J. Henning (Eds.), “Models, Simulations, and the Reduction of Complexity,” de Gruyter, 2013. ISBN: 978-3110313604
- [32] J.d’A. Guedes, S.A. Crabtree, R.K. Bocinsky, and T.A. Kohler, “21st-Century Approaches to Ancient Problems: Climate and Society,” *PNAS*, vol. 113, no.51, pp. 14483-14491. <https://doi.org/10.1073/pnas.1616188113>
- [33] M. Llobera, “Archaeological Visualization: Towards an Archaeological Information Science (AISC),” *J Archaeol Method Theory*, vol. 18, no. 3, pp. 193-223, September 2011. <https://doi.org/10.1007/s10816-010-9098-4>
- [34] A.J. Clark, “Seeing Beneath the Soil, Prospecting Methods in Archaeology,” Batsford London, 1990. ISBN: 978-0415214407
- [35] A. Collar, F. Coward, T. Brughmans, and B.J. Mills, “Networks in Archaeology: Phenomena, Abstraction, Representation,” *J Archaeol Method Theory*, vol. 22, no. 1, pp. 1–32. January 2015. <https://doi.org/10.1007/s10816-014-9235-6>

QoS-Ensured Cooperative Vehicular Communications

Radwa Ahmed Osman^{*#}, and Xiao-Hong Peng^{*}

^{*}School of Engineering & Applied Science, Aston University
Birmingham, UK

[#]College of Engineering & Technology, Arab Academy for Science, Technology & Maritime Transport
Alexandria, Egypt

email: {osmanmra, x-h.peng}@aston.ac.uk, radwa.ahmed@aast.edu

Abstract— Vehicular communications aims to provide effective and sustainable connections between vehicles and between vehicles and infrastructure. One of the key challenges in vehicular communications, especially vehicular-to-infrastructure (V2I), is to safeguard the best performance at the lowest resource cost. In this paper, we propose a quality-of-service (QoS) ensured V2I approach which is supported by cooperative communications via vehicle-to-vehicle (V2V) relaying, to maximize throughput and minimize energy consumption among different transmission schemes. Based on the closed-form expressions of the outage probability, throughput, energy efficiency and packet loss rate for various transmission schemes concerned, we demonstrate the performance and optimization strategy of the proposed approach. We also show how the best performance trade-off between system reliability and efficiency under various environmental conditions can be achieved.

Keywords— QoS; V2I; V2V; cooperative communications.

I. INTRODUCTION

Vehicular communications in the form of vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and their combinations called V2X are becoming one of key technologies to support connected and autonomous vehicles. They are also essential for enabling diverse applications associated with traffic safety, operation efficiency and infotainment [1] [2] [3]. In a vehicular network, road users are able to access Internet services such as traffic condition broadcast, video streaming, digital map downloading, and information of road hazard and accident alarm, via fixed roadside units through V2X communications. The most recent research in this area has been focused on the vehicular ad-hoc network (VANET) [4] [5], including its connection to the Fourth-Generation or Long-Term Evolution (LTE and LTE-Advance) cellular networks and the provision of the required solutions for achieving low latency and high reliability in vehicular communications [6].

IEEE 802.11p is considered one of the popular standards designed for vehicular communications, but it has showed obvious drawbacks such as low reliability, unbounded delay, hidden node problem and intermittent V2I connectivity [7].

To tackle the problem encountered when improving quality of service (QoS), cooperative communications techniques can be applied to enhance transmission reliability by creating diversity [8]. In this case, mobile nodes (vehicles) can help each other through relaying other node's data and sharing their limited resources to improve loss performance and increase transmission coverage. However, the performance enhancement by using relays nodes is constrained by the power (energy) budget imposed and high mobility in the vehicular network [9]. This issue can potentially impede the delivery of QoS in the V2I approach. In addition, routing in vehicular networks through cooperative relaying plays an important role in forwarding the required data to other vehicles with enhanced system performance [10].

In this work, we examine the performance of both cooperative and non-cooperative transmission schemes in the context of a vehicular network, including energy consumption, throughput and packet loss rate under different conditions, such as transmission distance, relaying method and channel condition (path loss exponent). We also identify optimal transmission schemes for the whole network in a changing environment. To the best of our knowledge, the proposed approach is unique since it provides an efficient way to find the best method for transmission between any V2I links with the help of V2V and based on the models we derive.

The remainder of this paper is organized as follows. Section II discusses the relevance of this research with other work. The system models for both cooperative and non-cooperative transmission schemes for V2I communications are presented in Section III. Simulation results produced by Matlab and NS-2 and discussions are presented in Section IV. Finally, the paper is concluded in Section V.

II. RELATED WORK

Cooperative communications technologies for VANETs have been studied extensively, where two of the most common protocols of this technology are Amplify-and-Forward (AF) and Decode-and-Forward (DF) [11]. Cooperative or polarization diversity is implemented by applying these protocols to exploit the broadcast nature of

wireless channels and use relays to improve link reliability and throughput [12]. In addition, the use of graph theory to formulate the problem of cooperative communications scheduling in vehicular networks is proposed in [13], in order to improve the throughput and spectral efficiency of the network.

Enhancing system efficiency is a key issue in applying cooperative communications in V2I approaches, depending on the connectivity probabilities in V2I and V2V communication scenarios in one-way and two-way platooning based VANETs [14]. Smart Antenna technologies can also contribute to the increment of the service coverage and system throughput of V2I [15]. The capacity of V2I communications can be maximized by an iterative resource allocation method [16] and the efficiency of V2I communications can be improved by applying a scheme called Distributed Sorting Mechanism (DSM) [17]. To improve power efficiency in V2I communication networks, [18] proposed a joint power and sub-carrier assignment policy under delay-aware QoS requirements. In addition, the strong dependence on the environment due to multipath propagation is also presented for a distributed energy-efficient routing method [19].

Furthermore, [20] proposed an adaptive rate adaptation algorithm integrated with a power control scheme. It minimizes energy consumption by appropriately adjusting vehicle's transmitting power, reducing network congestion and improving collision avoidance in vehicular networks. In [21] a sub-channel power control algorithm is proposed and the associated optimization problem is formulated to handle increased co-channel interference due to high mobility of vehicles in the network.

Although there have been different methods reported for improving performances in vehicular networks, there is a lack of information regarding how to choose specific transmission schemes that can ensure the best QoS under different conditions in terms of the number of relaying branches and the number of relays for a given distance between source and destination nodes.

In this paper, based on the initial work proposed in [1] our focus will be the identification of the conditions for establishing appropriate transmission strategies among different commonly used transmission schemes, including both cooperative and non-cooperative schemes for V2I communications. In cooperative communications nodes/vehicles not only transmit their own information, but also relay other nodes' information to a common destination. On the other hand, in non-cooperative communications nodes send their information directly to the destination, without relaying for one another. Our approach is to utilize the analytical models derived for these transmission schemes and to evaluate their performances in reliability, energy efficiency and throughput. Based on the trade-offs between cooperative and non-cooperative transmission schemes, we will show how to achieve the best performance through adaptive cooperative communications.

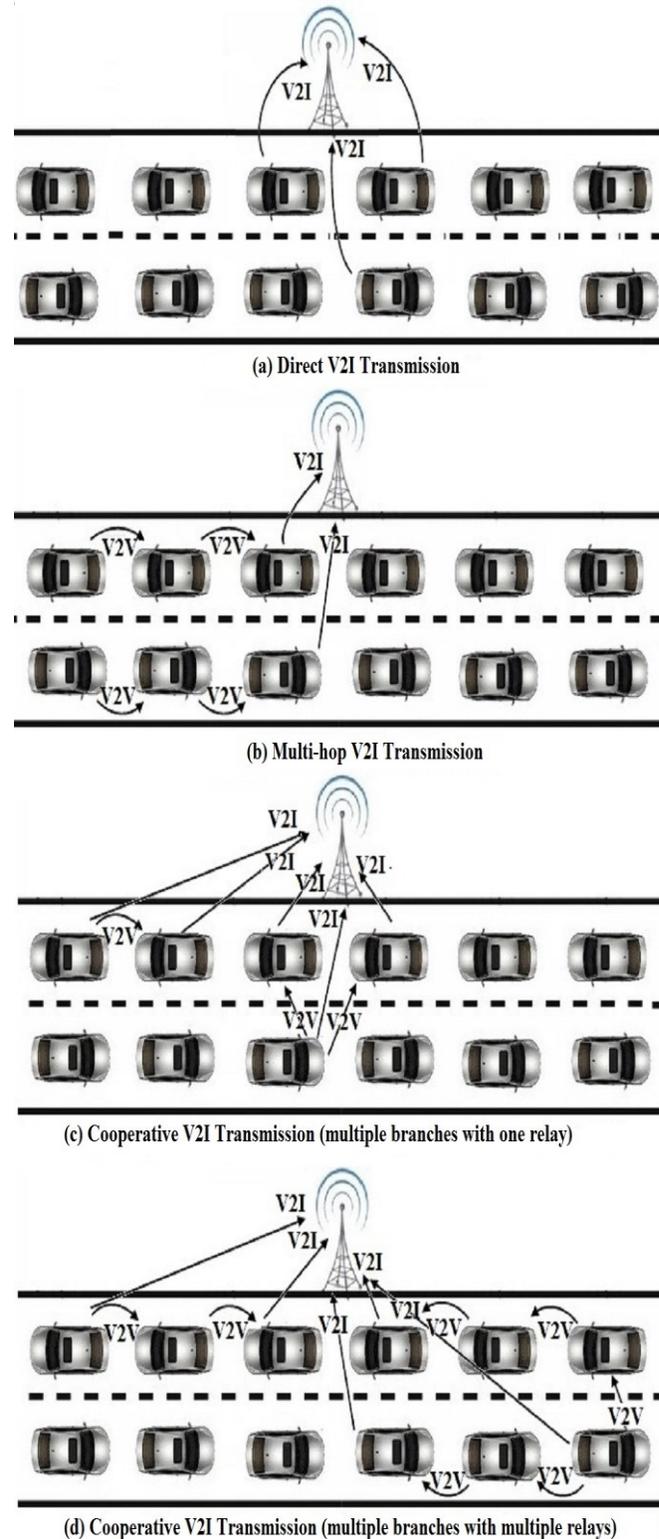


Figure 1. Different V2I transmission schemes.

III. SYSTEM MODEL

In this section, the analytical models of the required transmitting power, outage probability, energy consumption, throughput and packet loss rate in the context of a V2I network are established for both cooperative and non-cooperative transmission schemes. Based on these models, an adaptive transmission strategy can be developed to optimize the system performance.

Given a V2X network with L vehicles, for any vehicle-to-infrastructure pair (V, I) , where $V \in \{1, \dots, L\}$, the goal of the optimization proposed in this work in connection to QoS is achieved by either minimizing the total energy consumed per bit (or energy efficiency) given an outage probability target, or maximizing the end-to-end throughput, (or minimizing the packet loss rate) based on the transmission distance between V2I pairs, i.e.:

$$\begin{aligned} \text{Min } \sum E_{bi} & \quad \text{s.t. } \{p_{outVI}\} \text{ or} \\ \text{Max } \sum S_{thi} & \quad \text{s.t. } \{d_{VI}\} \end{aligned} \quad (1)$$

where E_{bi} and S_{thi} are the energy consumed per bit and throughput, respectively, of the i -th path between a vehicle (V) and infrastructure (I), p_{outVI} and d_{VI} are the fixed outage probability target and the total transmission distance between V and I . As V2V is part of the overall V2I, so the outage probability given in (1) is for the end-to-end V2I route which includes initial V2V links.

Four transmission schemes in the context of V2I are identified in Figure 1, including single-hop direct V2I (1a), multi-hop V2I via V2V (1b), cooperative V2I with a single relay in each relaying branch (1c), and cooperative V2I with multiple relays in each relaying branch (1d). In this work, we intend to examine the performances of different transmission schemes in terms of energy efficiency, throughput and packet loss rate, and to optimize them under different environmental conditions. In these schemes, the transmission path that forms a V2V link is selected based on the distance measurement and channel conditions in terms of the path loss exponent of the V2V link.

We consider a V2I network in which transmission links are subject to narrowband Rayleigh fading with additive white Gaussian noise (AWGN) and propagation path-loss. The channel fades for different links are assumed to be statistically mutually independent. For medium access, vehicle nodes are assumed to transmit over orthogonal channels through using the service channels specified in IEEE 801.11p [5], thus no mutual interference is considered in this system model. These channels can be reused by other vehicle away from a certain distance.

A. Non-Cooperative Transmission Scheme

Consider the transmission scheme for a direct link (V, I) as shown in Figure 1a where no relaying paths are involved. We use P_{SDir} to denote the source transmission power for this case. For direct transmissions in the V - I link, the received symbol r_{VI} and the spectral efficiency R_s (bits/sec/Hz) can be modeled as [22] [23]:

$$r_{VI} = \sqrt{P_{SDir} d_{VI}^{-\alpha}} h_{VI} s + n_{SD} \quad (2)$$

$$R_s = \frac{1}{2} \log_2(1 + SNR_{VI}) \quad (3)$$

where d_{VI} is the distance and h_{VI} is the channel coefficient of the V - I link, α is the path loss exponent, s is the transmitted symbol with unit power, and n_{SD} represents the AWGN noise vector, with variance $N_o/2$ where N_o is the thermal noise power spectral density (W/Hz).

The log-normal environment shadowing path loss model at a distance d_{ij} between node i and node j is given by [24]:

$$\gamma_{ij} [dB] = PL(d_o) + 10\alpha \log_{10}\left(\frac{d_{ij}}{d_o}\right) + X_\sigma \quad (4)$$

where X_σ is a zero-mean Gaussian distributed random variable with standard deviation σ and with some time correlation. This variable is zero if no shadowing effect exists. The $PL(d_o)$ is the path loss at a reference distance d_o in dB. The Signal-to-Noise Ratio (SNR) of the V - I link is expressed as [22]:

$$SNR_{VI} = \frac{P_{SDir} |h_{VI}|^2 \gamma_{VI}}{N} \quad (5)$$

where $N = N_o B$ is the noise power, and B is the system bandwidth in Hertz.

An outage occurs when the SNR at the receiver falls below a threshold β which allows error free decoding. This threshold is defined as $\beta = 2^{2R_s} - 1$, where R_s is the required system spectral efficiency. The outage probability of the single-hop transmission is given by [22] [23]:

$$p_{outVI} = p(SNR_{VI} \leq \beta) = 1 - e^{-\frac{(2^{2R_s} - 1)N}{P_{SDir} |h_{VI}|^2 \gamma_{VI}}} \quad (6)$$

Energy consumption is largely proportional to the requirement of maintaining a certain level of transmission reliability or the successful transmission rate. In order to maintain a required level of reliability, denoted by U , which is related to the reliability of a transmission link, the minimum outage probability is defined as:

$$p_{outDir} \leq 1 - U \quad (7)$$

Combining (6) and (7) and taking the nature logarithm on both sides of the expression, we have:

$$\frac{(2^{2R_s} - 1)N}{P_{SDir} |h_{VI}|^2 \gamma_{VI}} \leq \ln(U^{-1}) \quad (8)$$

The main objective for the performance optimization of a V2I network is to minimize the total energy consumption under different environmental conditions. Thus, the transmit power required to satisfy the reliability requirement or constrained by the outage probability for the direct transmission must be:

$$P_{SDir} \geq (2^{2R_s} - 1) \frac{N}{|h_{VI}|^2 \gamma_{VI}} (\ln(U^{-1}))^{-1} \quad (9)$$

Therefore, the total consumed energy per bit (J/bit) for the direct transmission mode can be expressed as:

$$E_{bDir} = \frac{P_{AM,Dir} + P_C}{R_b} \quad (10)$$

where

$$P_C = P_{Tx} + P_{Rx}$$

$$P_{AM,Dir} = \frac{\xi}{\eta} P_{SDir} \quad (11)$$

where $P_{AM,Dir}$ is the power amplifier consumption for direct transmission which depends on the drain efficiency of the amplifier η , the average peak-to-peak ratio ξ , and the transmit power P_{SDir} , $R_b = R_s B$ is the data rate in bits/s, B is the system bandwidth, and P_C is the power consumed by the internal circuitry for transmitting (P_{Tx}) and receiving (P_{Rx}).

The throughput S_{th} and packet loss rate PLR can be simply defined, i.e.:

$$S_{th} = \frac{\text{Total Received Payload}}{\text{Total Transmitted Time}} \quad (12)$$

$$PLR = \frac{\text{Total Sent Packets} - \text{Total Received Packets}}{\text{Total Sent Packets}} \quad (13)$$

The multi-hop non-cooperative transmission scheme with n ($n \geq 1$) relays is shown in Figure 1b. Each relay is able to detect whether or not the packet was received correctly and will forward the information to the destination only in the case of the packet being correctly received. Otherwise, the packet is considered lost.

Given the outage probabilities of individual hops, i.e., p_{outVR_1} (from a vehicle to relay 1), $p_{outR_1R_2}$ (from relay 1 to relay 2), ..., p_{outR_nI} (from relay n to infrastructure), the outage probability of the multi-hop link, p_{outMH} , is given by:

$$p_{outMH} = 1 - (1 - p_{outVR_1})(1 - p_{outR_1R_2}) \dots (1 - p_{outR_nI}) \quad (14)$$

With the same mathematical treatment as in (6), p_{outMH} becomes:

$$p_{outMH} = 1 - e^{-(2^{2R_s} - 1)N} y \quad (15)$$

$$\text{where } y = \left(\frac{1}{P_{Vr1} |h_{Vr1}|^2 \gamma_{Vr1}} + \sum_{i=2}^n \frac{1}{P_{r1ri} |h_{r1ri}|^2 \gamma_{r1ri}} + \frac{1}{P_{mI} |h_{mI}|^2 \gamma_{mI}} \right)$$

We set the transmit power to be proportional to the distance between two communicating nodes. For broadcast transmission, e.g., when the source transmits, the longest distance, i.e., the distance between the source and the destination d_{SD} , is considered. Hence, the power between two communicating nodes is given by:

$$P_{ij} = X P_{SD} \quad (16)$$

where X denotes the power coefficient between node i and node j . In our model, we assume that the value of X depends on the distance of the vehicle-infrastructure (VI), relay-relay (RR) or relay-infrastructure (RI) link. For example, the transmit power for the relay- infrastructure link is:

$$P_{RI} = (\lambda_{RI})^\alpha P_{VI} \quad (17)$$

$$\text{where } \lambda_{RI} = \frac{d_{RI}}{d_{VI}}$$

The power minimization problem is specified in a similar way to (7), i.e.:

$$p_{outMH} \leq 1 - U \quad (18)$$

and the power P_{SMH} is bounded by:

$$P_{SMH} \geq (2^{2R_s} - 1) N y (\ln(U^{-1}))^{-1} \quad (19)$$

Then, the total consumed energy per bit and the total consumed power for the multi-hop direct transmission are expressed as:

$$E_{bMH} = (p_{outMH}) \frac{P_{AM,MH} + P_C}{R_b} + (1 - p_{outMH}) \frac{(n * X + 1) P_{AM,MH} + (n + 1) P_C}{R_b} \quad (20)$$

$$P_{totMH} = (p_{outMH}) (P_{AM,MH} + P_C) + (1 - p_{outMH}) ((n * X + 1) P_{AM,MH} + (n + 1) P_C) \quad (21)$$

where $P_{AM,MH}$ is the power amplifier consumption for multi-hop transmission.

The first term on the right-hand side of (20) corresponds to the consumed energy when the relay is not able to correctly decode the message from the vehicle, which means that this link is in outage. In this case, only the source vehicle

consumes transmitting power, and the destination node and K relays consume receiving power. The second term counts for the event that the V - I link is not in outage, hence the relay's transmitting and processing power, and the extra receiving power at the infrastructure are involved.

B. Cooperative Transmission Scheme

In cooperative transmission, the sender V broadcasts its symbol to all potential receivers including the destination I and relays in the current time slot. The received symbol by relays, r_{sr} , the received symbol by the destination from relays, r_{rd} , and the spectral efficiency R_S can be expressed as:

$$r_{Vr} = \sqrt{P_S d_{Vr}^{-\alpha}} h_{Vr} s + n_{Vr} \quad (22)$$

$$r_{rI} = \sqrt{P_C d_{rI}^{-\alpha}} h_{rI} s + n_{rI} \quad (23)$$

$$R_s = \frac{1}{2} \log_2(1 + SNR_{Vr} + SNR_{rI}) \quad (24)$$

where P_S is the transmitted power of the source and P_C is the transmitted power of relays, h_{Vr} and h_{rI} are the channel coefficients of the vehicle-relay link and the relay-infrastructure link, respectively, n_{Vr} and n_{rI} are the AWGN noise vectors of the vehicle-relay link and the relay-infrastructure link, respectively. SNR_{Vr} and SNR_{rI} are the signal-to-noise ratios of the V - r link and r - I link, respectively.

Two types of cooperative transmission schemes are considered here: 1) using multiple cooperative relaying branches with a single relay in each branch (MBSR) (Figure 1c), and 2) multiple relaying branches with multiple relays in each branch (MBMR) (Figure 1d). The selective decode-and-forward (SDF) relaying protocol is used in these two schemes and relays perform cooperation when the information from the source is correctly received by them. For the transmission scheme shown in Figure 1c, the outage probability is given by jointly considering the outages in V - I , V - R and R - I links, i.e.:

$$P_{outMB} = P_{outVI} (P_{outVr} + P_{outrI} - P_{outVr} P_{outrI})^K \quad (25)$$

Based on the derivation methods used in Section III-A, the following close-form expressions can be readily obtained.

1): The outage probability of cooperative transmission with multiple (K) branches with each having multiple relays (n):

$$P_{outMHB} \approx (2^{2R_s} - 1)^{K+1} N^{K+1} C \quad (26)$$

where

$$C = \left(\frac{1}{|h_{Vr1}|^2 \gamma_{Vr1}} + \sum_{i=2}^n \frac{1}{\lambda_{ri-1ri}^{-\alpha} |h_{ri-1ri}|^2 \gamma_{ri-1ri}} + \frac{1}{\lambda_{rml}^{-\alpha} |h_{rml}|^2 \gamma_{rml}} \right)^K$$

2): The power minimization problem is specified in a similar way to (7):

$$P_{outMHB} \leq 1 - U \quad (27)$$

3): The lower bound of power for cooperative transmission with multiple (K) branches and multiple relays (n):

$$P_{MHB} \geq (2^{2R_s} - 1) N \left(\frac{1}{|h_{Vr1}|^2 \gamma_{Vr1}} C \right)^{\left(\frac{1}{K+1}\right)} (\ln(U^{-1}))^{\left(\frac{1}{K+1}\right)} \quad (28)$$

4): The total consumed energy per bit and the total consumed power for this cooperative transmission scheme:

$$E_{bMHB} = (P_{outVr}) \frac{P_{AM,MHB} + P_{Tx} + (K+1)P_{Rx}}{R_b} \quad (29)$$

$$+ (1 - P_{outVr}) \frac{(K * n * X + 1)P_{AM,MHB} + (K * n + 1)P_{Tx} + (K * n + 2)P_{Rx}}{R_b}$$

$$P_{totMHB} = (P_{outVr}) (P_{AM,MHB} + P_{Tx} + (K+1)P_{Rx}) \quad (30)$$

$$+ (1 - P_{outVr}) ((K * n * X + 1)P_{AM,MHB} + (K * n + 1)P_{Tx} + (K * n + 2)P_{Rx})$$

The transmit power at relays can be reduced and consequently the energy efficiency will be improved by implementing the cooperative communications schemes, which are particularly suitable for long-range transmissions. The related results will be shown in Section IV.

IV. NUMERICAL RESULTS AND DISCUSSION

In this section, we examine the performances of different transmission schemes through Matlab and NS-2 simulations, in terms of energy efficiency (energy consumption per bit), throughput, packet loss rate, and optimum required number of branches and relays for different transmission distances in V2I links. We then reveal the conditions for selecting the optimal transmission schemes through comparisons between them. The network settings used for simulation are listed in TABLE I. Assume the spectral efficiency R_s in this scenario to be 2 bits/sec/Hz, and the required system reliability level to be 0.999. To generate mobility, related mobility-files are created in NS-2 simulation. In addition, we assume that all the vehicles are running at the same speed and keeping the same distance with each other.

In Figure 2, the energy performances of both cooperative and non-cooperative schemes are illustrated and compared. As we can see, the non-cooperative direct transmission has the lowest energy cost than all others transmission schemes for short-range ($d_{VI} < 33$ m); the non-cooperative transmission using multi-hop relays outperforms the direct transmission for the range $33\text{m} < d_{VI} < 43$ m and, in particular, transmission using two intermediate relays ($n=2$) nodes has the lowest energy consumption for this range.

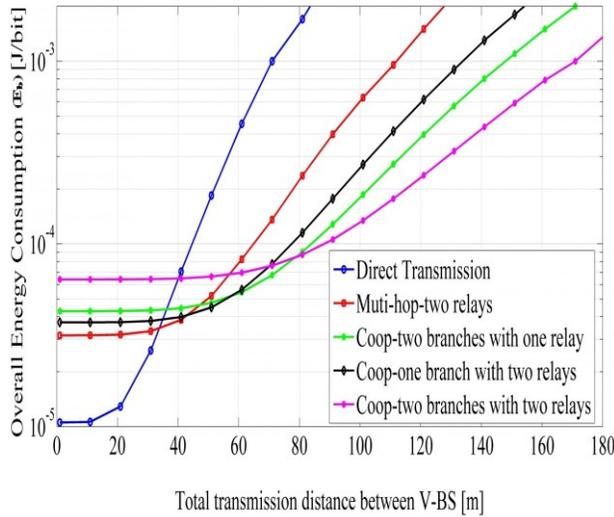


Figure 2. Total energy consumed vs total transmitted distance.

The cooperative transmission outperforms the non-cooperative transmission schemes for the range $43m < d_{Vl} < 58m$, and the transmission using one branch with two relays ($K=1, n=2$) has the lowest energy consumption for this range. As distance continuously increases, the lowest energy consumption is achieved by transmission using two branches with one relay ($K=2, n=1$) for $58 < d_{Vl} < 80m$, and by transmission using two branches with two relays ($K=2, n=2$) for $d_{Vl} > 80m$, respectively.

As shown in Figure 3, the non-cooperative direct transmission has much higher energy consumption than the optimum transmission scheme which is chosen based on the transmission distance between vehicles and infrastructure.

The results in Figure 4 show the energy performance for multi-branch and multi-relay scenarios for five different transmission distances in V2I links, i.e., 20m, 40m, 50m, 70m and 90m. Under each distance the energy performance is examined against the number of relays (n) for a different number of branches $K=1, 2, \dots, 5$, employed by cooperative communications. From this examination, the optimal number of relays per branch can be found among different scenarios.

When $d_{Vl}=20m$, the direct transmission scheme which does not need any diverse branches can be the most energy efficient transmission scheme as shown in Figure 4a. However, when increasing the distance a clear trend is shown that cooperative transmission schemes are becoming more energy efficient than the direct transmission scheme for a certain range of the number of relays used per branch. And this range is widening as the distance increases. As we can see, cooperative transmission can outperform direct transmission for $n < 2$ when $d_{Vl}=40m$, for $n < 3$ when $d_{Vl}=50m$, for $n < 4$ when $d_{Vl}=70m$, and for $n < 5$ when $d_{Vl}=90m$, as shown in Figures 4b-4e, respectively.

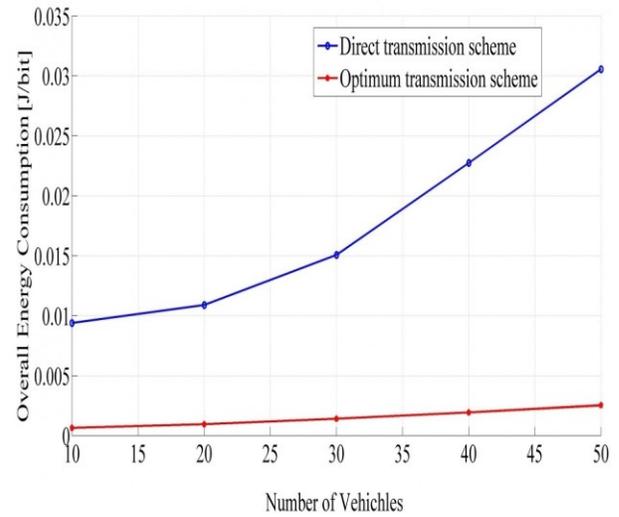


Figure 3. Overall energy consumption vs number of vehicles.

TABLE I. SIMULATION PARAMETERS

Parameters	Value
N_0	-174 dBm
B	10 kHz
R_s	2 bits/sec/Hz [17].
P_{TX}	97.9 mW [17]
P_{RX}	112.2 mW [17]
η	0.35
ξ	0.5
Packet Size	512 bytes
f_c	5.9 GHz
α	3
Simulation time	1000 sec
Nodes	10/20/30/40/50
Velocity	5 km/h, 20 km/h, 60 km/h
Traffic Agent	TCP
Mac Protocol	IEEE 802.11p
Queue	PriQueue with size of 50 Packets
Propagation model	Log-normal shadowing Model (LOS)
Antenna	Omni-directional with height of 1m
Routing Protocol	AODV
Number of Seed	3

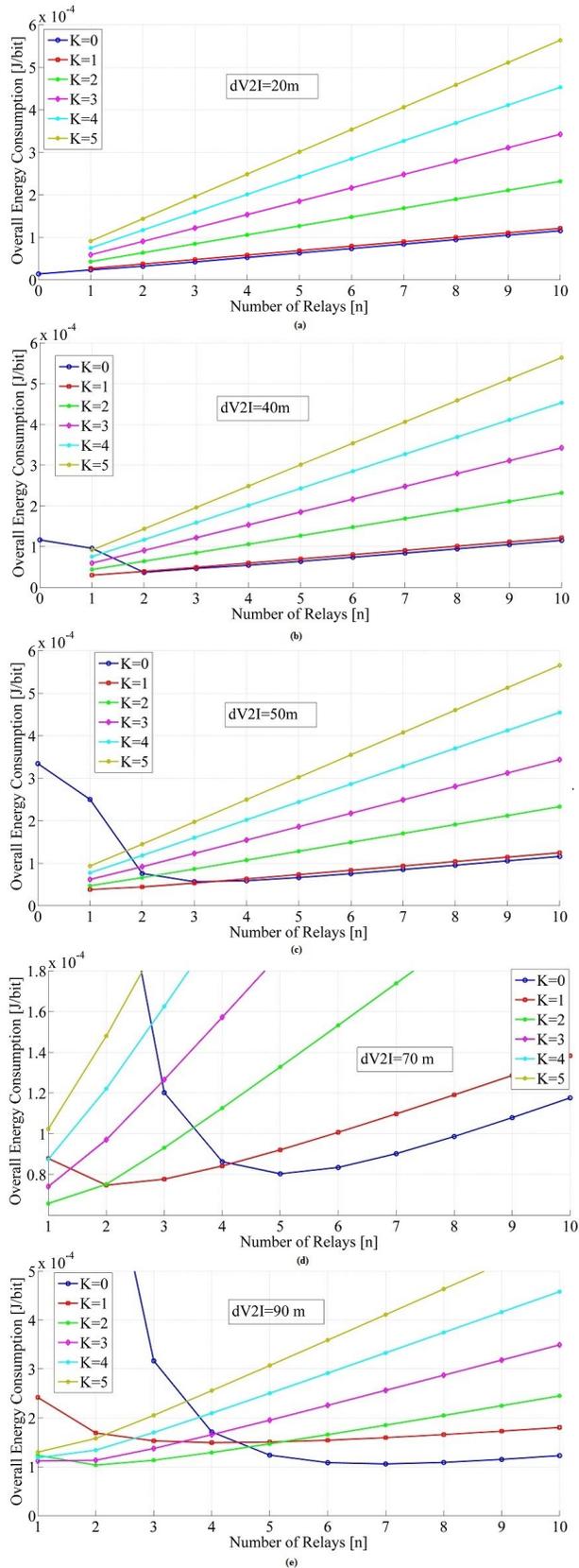


Figure 4. Overall energy consumption vs number of relays.

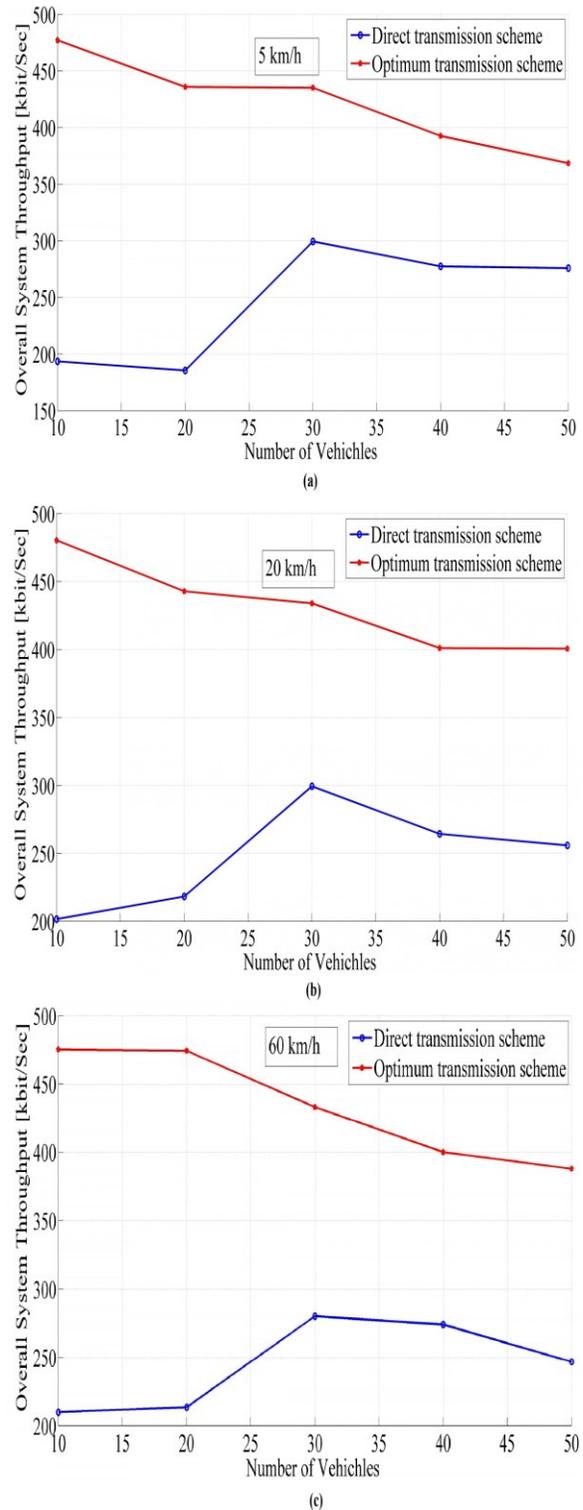


Figure 5. Overall system throughput vs number of vehicles.

The optimal transmission scheme for energy performance will be determined based on the transmission method used

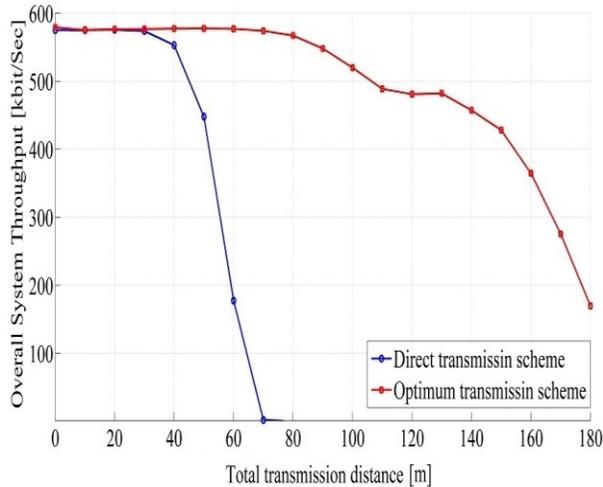


Figure 6. Overall system throughput vs total transmission distance.

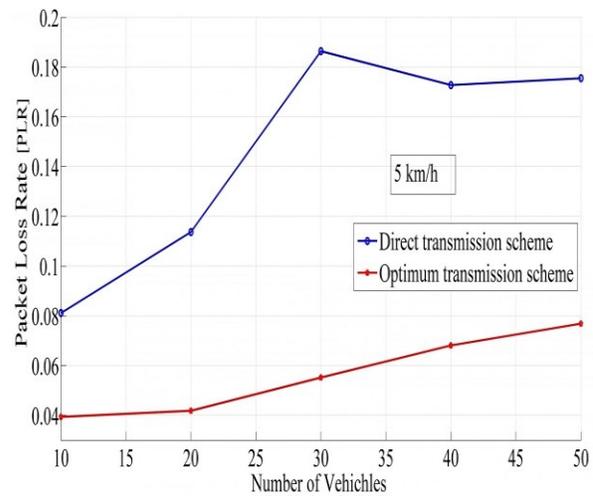
and conditions discussed above. For instance, the cooperative transmission scheme with one branch ($K=1$) is optimal for $n=2, 3$ and 4 when $d_{V2I}=70m$, while the cooperative transmission scheme with two branches ($K=2$) is optimal for $n=2, 3$ and 4 when $d_{V2I}=90m$. This indicates that an adaptive strategy can be applied to select the transmission scheme dynamically so that the best performance can be achieved and remained.

The overall system throughput is shown in Figure 5 for three different vehicle velocities. The optimum transmission schemes through cooperative communications clearly outperform the direct transmission schemes in all cases due to the impact of diversity created by cooperative transmission. It is also noticed that the throughput of the optimum transmission scheme decreases when the number of transmitting vehicles increases. This is mainly due to congestion in medium access and increased operation overhead at the nodes that have dual responsibilities as the source as well as the relay.

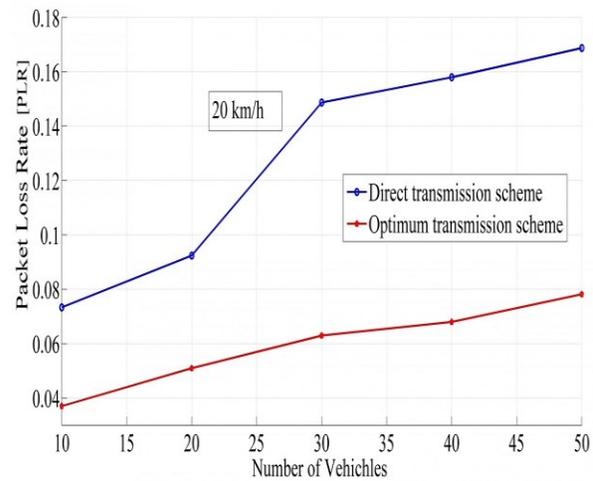
Again, the overall system throughput is examined in Figure 6 but against the total transmission distance of a V2I link for both direct and optimum transmission schemes. The optimum transmission schemes clearly outperform the direct transmission schemes for all transmission distances. It is shown that the throughput of the direct transmission scheme decreases dramatically compared with the optimum transmission scheme when the distance of the V2I link involved exceeds 30 m.

Figure 7 depicts the overall packet loss rate for direct transmission and optimum transmission schemes versus the number of transmitting vehicles for different vehicle velocities. As it is shown, the packet loss rate increases when the number of transmitted vehicles increases for all the transmission schemes, which is caused by network congestion and correlated with the corresponding performance in throughput as shown in Figure 5.

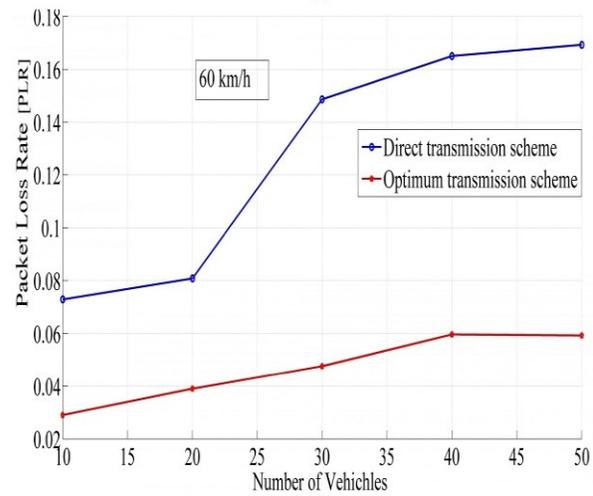
The overall system packet loss rates for direct transmission and optimum transmission schemes for each transmission distance are illustrated in Figure 8. The optimum transmission scheme clearly performs better than the direct



(a)



(b)



(c)

Figure 7. Packet loss rate vs number of vehicles.

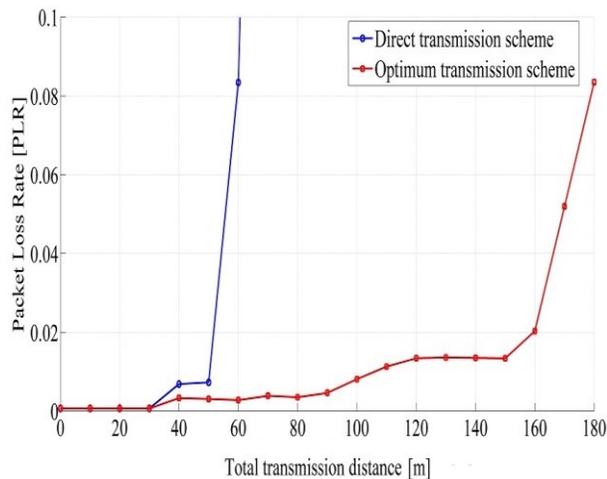


Figure 8. Packet loss rate vs total transmission distance.

transmission scheme for all transmission distances. These results are also correlated with those in Figure 6 for throughput performance, i.e., the dramatic reduction in throughput when the distance exceeds 30m is mainly caused by a sharp increase in the packet loss rate when the distance increases. It is worth mentioning that the optimum transmission schemes have much lower packet loss rates than the direct transmission schemes since when relays are used the transmission distances between adjacent nodes are reduced and, at the same time, the transmission reliability is improved due to the diversity generated through cooperative communications.

Due to the network settings in our work where most vehicles have a fairly large distance between them and the roadside base station, no major difference in performance is observed when increasing the velocity of vehicles, as shown in Figures 5 and 7. In contrast, as discussed above, the performance such as throughput is correlated with the number of vehicles which are connected to the same base station.

There are a number of factors affecting performances such as energy consumption, throughput and packet loss rate in V2X networks. Cooperative transmission utilizes additional paths and intermediate nodes to create diversity, which may cost more energy. However, this can be compensated by the diversity generated that can lower the probability of link failure and consequently reduce the number of retransmissions. Diversity can also be enhanced with the increased number of relaying branches, but this increase could be marginal when the number of branches is large and these branches are not all uncorrelated in this case.

Regarding improving the QoS performance, a clear advantage of cooperative transmission over direct transmission has been demonstrated in our results. The packet loss rate of cooperative transmission becomes much lower than that of direct transmission when the number of vehicles increases as shown in Figure 7. As a result, the throughput performance of cooperative transmission can always outperform direct transmission for a wide range of the vehicle density, as shown in Figure 4. It is also noticed that the direct

transmission schemes can also perform better than cooperative transmission under certain circumstances, as discussed above and shown in Figures 2 and 4.

To achieve the best energy performance for a specified application, proper transmission schemes should be selected in an environment affected by varying environmental conditions, such as overall distance, d_{SD} ; and channel quality in terms of α . The findings of this work can assist deciding when and how the cooperative or non-cooperative transmission scheme should be employed. Based on our investigation, an energy-efficient or throughput-centric transmission strategy can be formed in a V2X network by adaptively choosing proper transmission schemes under different network and transmission conditions. This involves determining the number of relaying branches and the number of relays if the cooperative scheme is to be used. By doing so, energy saving could be significant even with the direct transmission scheme in certain conditions, as demonstrated by our results.

V. CONCLUSION

We have investigated different transmission schemes for their performances on energy efficiency, throughput and packet loss rate in a vehicular network. Based on the models derived for outage probability, energy efficiency, throughput and packet loss rate, we have shown that both cooperative and non-cooperative transmission schemes can exhibit the best performance under certain environmental conditions. In addition, we have shown the required optimum numbers of branches and relays in each branch in order to enhance the system performance. The optimal transmission scheme can be identified given the distance between the source and destination nodes in a V2X network. The results presented in this paper can be used to form an adaptive transmission strategy that is able to select appropriate transmission schemes in a changing environment to maintain the best QoS performance in a dynamic way, in terms of achieving the highest throughput with a fixed energy budget or the lowest energy cost for a given throughput target.

REFERENCES

- [1] R. A. Osman, X.-H. Peng, and M. A. Omar, "Enhancing QoS in Vehicle-to-Infrastructure Approaches through Adaptive Cooperative Communications," The Sixth International Conference on Advances in Vehicular Systems (VEHICULAR), pp. 13-18, 2017.
- [2] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil, "Vehicular networking: a survey and tutorial on requirements, architectures, challenges, standards and solutions," IEEE Communications Surveys & Tutorials, vol. 13, pp. 584-616, July 2011.
- [3] E. C. Eze, S.-J. Zhang, E.-J. Liu, and J. C. Eze, "Advances in vehicular ad-hoc networks (VANETS): Challenges and road-map for future development" International Journal of Automation and Computing, vol. 13, pp. 1-18, February 2016.
- [4] S. Wu, C. Chen, and M. Chen, "An asymmetric and asynchronous energy conservation protocol for vehicular networks," IEEE Trans. Mobile Computing, vol. 9, pp. 98-111, January 2010.

- [5] M. J. Booyens, S. Zeadally, and G.-J. V. Rooyen, "Survey of media access control protocols for vehicular ad hoc networks," *IET Communications*, vol. 5, pp. 1619-1631, March 2011.
- [6] G. el mouna Zhioua, N. Tabbane, H. Labiod, and S. Tabbane, "A fuzzy multi-metric QoS-balancing gateway selection algorithm in a clustered VANET to LTE-advanced hybrid cellular network," *IEEE Trans. Vehicular Technology*, vol. 64, pp. 804-817, February 2015.
- [7] M. Amadeo, C. Campolo and A. Molinaro, "Enhancing IEEE 802.11p/WAVE to provide infotainment applications in VANETs," *Ad Hoc Networks*, vol. 10, no. 2, pp. 253-269, March 2012.
- [8] A. Nasri, R. Schober, and I. F. Blake, "Performance and optimization of amplify-and-forward cooperative diversity systems in generic noise and interference," *IEEE Trans. Wireless Communications*, vol. 10, pp. 1132-1143, January 2011.
- [9] D. Xie, W. Wei, Y. Wang, and H. Zhu, "Tradeoff between throughput and energy consumption in multirate wireless sensor networks," *IEEE Sensor Journal*, vol. 12, pp. 3667-3676, August 2013.
- [10] V. Duduku, V. A. Chekima, F. Wong, and J. A. Dargham, "A Study on Vehicular Ad Hoc Networks," 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS), 2-4 December 2015.
- [11] H. Hakim, H. Bouijemaa, and W. Ajib, "Single relay selection schemes for broadcast networks," *IEEE Trans. Wireless Communications*, vol. 12, pp. 2646-2657, May 2013.
- [12] S. Sohaib and D. K. C. So, "Asynchronous cooperative relaying for vehicle-to-vehicle communications," *IEEE Trans. Communications*, vol. 61, pp. 1732-1738, May 2013.
- [13] K. Zheng, F. Liu, Q. Zheng, W. Xiang, and W. Wang, "A graph-based cooperative scheduling scheme for vehicular networks," *IEEE Trans. Vehicular Technology*, vol. 62, pp. 1450-1458, May 2013.
- [14] C. Shao, S. Leng, Y. Zhang, A. Vinel, and M. Jonsson, "Performance analysis of connectivity probability and connectivity-aware MAC protocol design for platoon-based VANETs," *IEEE Trans. Vehicular Technology*, vol. 64, pp. 5596-5609, December 2015.
- [15] S.-Y. Pyun, H. Widiarti, Y.-J. Kwon, J. Son, and D. Cho, "Group-based channel access scheme for a V2I communications system using smart antenna," *IEEE Communications Letters*, vol. 15, pp. 804-806, August 2011.
- [16] S.-Y. Pyun, W. Lee, and D.-H. Cho, "Resource allocation for vehicle-to-infrastructure communications using directional transmission," *IEEE Trans. Intelligent Transportation Systems*, vol. 17, pp. 1183-1188, April 2016.
- [17] T.-Y. Wu, S. Guizani, W.-T. Lee, and K.-H. Liao, "Improving RSU service time by distributed sorting mechanism," *Ad Hoc Networks*, vol. 10, pp. 212-221, March 2012.
- [18] H. Zhang, Y. Ma, D. Yuan, and H.-H. Chen, "Quality-of-service driven power and sub-carrier allocation policy for vehicular communications networks," *IEEE Journal on Selected Area in Communications*, vol. 29, pp. 197-206, January 2011.
- [19] L. Azpilicueta, C. Vargas-Rosales, and F. Falcone, "Deterministic propagation prediction in transportation systems," *IEEE Vehicular Technology Magazine*, vol. 11, pp. 29-37, September 2016.
- [20] K. S. Nwizege, E. D. Nwiwure, and M. MacMammah, "Optimization of Power-Efficient Vehicular Networks in Rate Adaptation Algorithms," 16th International Conference on Computer Modelling and Simulation (UKSim-AMSS), 2014.
- [21] M. Patra, R. Thakur, and C. S. R. Murthy, "Improving delay and energy efficiency of vehicular networks using mobile femto access points," *IEEE Trans. Vehicular Technology*, vol. 66, pp. 1496-1505, May 2016.
- [22] G. G. de Oliveira Brante, M. T. Kakitani, and R. D. Souza, "Energy efficiency analysis of some cooperative and non-cooperative transmission schemes in wireless sensor networks," *IEEE Trans. Communications*, vol. 59, pp. 2671-2677, July 2011.
- [23] A. S. Ibrahim, Z. Han, and K.J.R. Liu, "Distributed energy-efficient routing in wireless networks," *IEEE Trans. Wireless Communications*, vol. 7, pp. 3930-3941, October 2008.
- [24] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

Measurement System Based on Arduino for Biogas Sensing: Development Considerations and Laboratory Scale Approach

Short Paper

Isaías González, Antonio José Calderón

Department of Electrical, Electronic and Automatic Engineering

University of Extremadura

Badajoz, Spain

e-mail: {ajcalde, igonzp}@unex.es

Abstract—Measurement systems are used in every scientific and technological process to measure and monitor a number of magnitudes. During last years, open source technology and, concretely, open source hardware devices are being introduced in different R&D projects due to their advantages of low cost, easy development, shared information in the Internet and so on. Arduino microcontroller is the most illustrative example of this kind of devices. This paper presents a measurement system based on Arduino board to perform biogas sensing. A set of sensors are responsible of measuring the gases that compose the biogas (hydrogen, methane, ammonia and carbon monoxide) and the results are displayed to the user via a small graphical monitoring interface. The development of the system is reported and initial results are provided to demonstrate the suitability of the microcontroller.

Keywords- measurement system; Arduino; open source; monitoring; biogas.

I. INTRODUCTION

Every scientific and technological process requires to measure and monitor variables of different nature (analog/digital) that inform about the behavior or state of such process. To this aim, measurement systems follow the common architecture depicted in Fig. 1. A physical process needs to be monitored and/or controlled, so a set of sensors measure the representative magnitudes. These sensors are linked to a microprocessor or microcontroller-based board, which is responsible of acquiring and processing their signals. In Fig. 1, both the sensors and the board are considered into the block named Measurement Acquisition System (MAS) for a higher level of generality. A Monitoring Interface (MI) exchanges data with the microcontroller in order to provide useful information, by means of numerical or graphical displays, to the user.

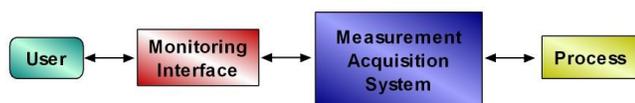


Figure 1. General diagram of measurement systems.

The present paper reports the development of a measuring system for biogas sensing using a development

board called Arduino, which is based on a microcontroller. This board is commonly referred to as Arduino microcontroller. Diverse considerations about the proposed system as well as initial results are also provided. Particularly, an Arduino Mega board and a set of four sensors (methane, ammonia, carbon monoxide and hydrogen) constitute the MAS; whereas a LCD screen plays the role of MI. This manuscript is an extended version of the congress contribution in [1].

A scientific domain where measurement systems are profusely applied is that related to Renewable Energy Sources (RES), which gains increasing attention both from governments and R&D initiatives around the world. As it well-known, the dependency on fossil fuels needs to be decreased so RES are progressively being introduced in the energetic scenario. Within RES, photovoltaic and wind energies are the most exploited due to their availability. Nonetheless, other important source is available as a result of the human activities, namely biogas. Biogas is a combustible fuel generated by the disposal of waste, and composed by a variable mixture of methane (50-70%) and carbon dioxide (30-50%), and also minimal quantities of nitrogen, hydrogen and water vapor (Fig. 2). Biogas can be produced in natural means or in a sealed, oxygen-free container named anaerobic digester, from the reactions of biodegradation of the organic matter, by the action of microorganisms in the absence of oxygen [2].

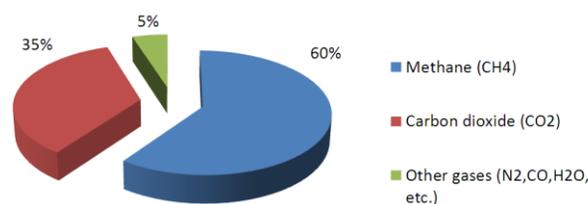


Figure 2. Composition of biogas.

Arduino microcontroller is nowadays the maximum representative of the open source hardware trend. Given its

versatility and widespread application for diverse scopes, in the authors' humble opinion, it is suitable to act as the core of a measurement system to determine the composition of biogas. In this sense, the main contribution of this work is the development of a fully functional measurement system for biogas using such an inexpensive open source microcontroller. This paper is envisioned to provide useful insights about the implementation of such system, which can be a valuable resource for scientists and scholars involved in R&D activities related to biogas fuel.

The remainder of the paper is as follows. Section II is devoted to provide a brief review of measuring systems and Arduino microcontroller applications in order to highlight the relevance of our proposal. Section III deals with the description of the developed system, both hardware and software approaches. Finally, main conclusions and further works are outlined in Section IV.

II. BRIEF REVIEW ABOUT MEASURING SYSTEMS AND ARDUINO

This section provides a background about the utilization of Arduino platforms, mainly in the context of measuring systems.

A noteworthy remark is that open source technology, comprising both hardware and software, constitutes an ever-growing trend during last years [3]. The open source tools provide the developer with the ability to perform in-depth configurations since the source code, schematics and documentation are available. As a consequence, these means offer serious advantages regarding low, or even inexistent, economic costs and custom-designed approaches.

On the contrary, the main drawback is that programming and/or hardware related skills are commonly required. In addition, there is no customer support, which can be an important boundary for professional solutions. This kind of tools is achieving increasing presence in different technological and scientific scenarios. For instance, modern paradigms like the global pervasive network of interconnected devices known as the Internet of Things (IoT) will take advantage of open source hardware and software projects [4, 5].

Regarding open source hardware, Raspberry Pi, BeagleBone, Phidget, OpenDAQ, Intel Edison and Arduino boards are devices of this nature that are being more and more exploited. Low-cost, ease of use and wide availability have opened a new door to democratize electronics [6]. Open source philosophy implies that a great variety of information, examples and tutorials are shared by a collaborative community through the Internet, embracing development based on this type of devices.

Arduino is a low-cost single-board microcontroller [7] and is the prevalent open source hardware device. Concerning R&D and Academia activities, Arduino constitutes a merging trend, which has become a powerful tool to develop different applications in the fields of data acquisition, automation and engineering in general [8].

The versatility of this microcontroller is validated by the numerous applications reported. Out of the measurement systems sphere, it has been used for humanoid robotic [9],

greenhouse control [10], fuzzy cognitive maps [11], remote laboratories for engineering education [12-15], ZigBee-based communication [16], cyber-physical systems [17, 18] and as electronic interface between Programmable Logic Controllers (PLCs) and sensors [19-21].

Concerning measurement-centered approaches, recent works report Arduino applications for agricultural instrumentation [22, 23], monitoring of photovoltaic systems [24-27], environmental monitoring [28, 29], Wireless Sensor Networks (WSNs) implementation [30], air quality monitoring [31], sensors deployment in Unmanned Aircraft Vehicles (UAVs) [32] or for monitoring hydrogen fuel cells [8, 33].

There is a wide diversity of Arduino boards with different sizes and resources. For instance, Uno, Yun, Mini, Nano, Mega, Duemilanove, Extreme, Lilypad, etc., so the developer can choose the model that fits better the application to implement. In the present case, the Mega 2560 R3 board has been chosen.

The developed system corresponds to the block diagram shown in Fig. 1. The functions of the MI as well as a further connectivity link are provided by the so-called shields. They are expansion cards with additional features that can be directly coupled on the main board. There are an increasing number of shields devoted to diverse tasks: data visualization, motor control, sensors connection, data storage, etc. Therefore, these cards facilitate the developments using Arduino, for instance, in the proposed approach; a display is used to provide information to the user as described in the next section.

Some advantages of the Arduino boards are now listed in order to emphasize the suitability of such a device for the purpose of the present work:

- Low cost. The cost of an Arduino Mega board is around 40€ whereas a commercial data acquisition board costs about 450€.
- Open hardware. All the schematics and documentation related to the Arduino microcontroller as well as to the shields are freely available in the Internet.
- Availability of expansion cards. The shields constitute an important advantage to promote the applications of Arduino since they facilitate the connection and configuration of additional features or devices.
- Great information shared by the community. Tutorials, forums and videos offer knowledge freely available that contribute to develop Arduino-based projects.
- Rapid development. As a result of the previous consideration, the time and effort required to design and implement systems including Arduino are shorten.
- Continuous improvements. The open source community continuously enhances the Arduino resources like libraries and shields, helping to

improve already existent systems or to design new approaches.

- Increasing application in R&D domain. As it has been expounded, Arduino is widely used as tool to develop a multiplicity of R&D applications.

III. DEVELOPMENT OF THE MEASUREMENT SYSTEM

This section deals with the description of the main features of the elements that compose the presented measurement system as well as their interconnection and configuration. For a better comprehension of the proposal, a block diagram of the whole system is provided in Fig. 3.

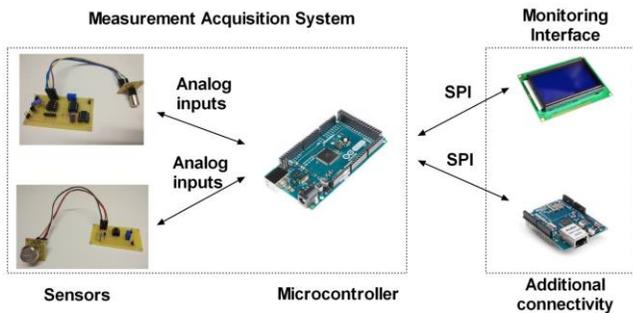


Figure 3. Diagram of the implemented prototype of measurement system.

A. Arduino microcontroller

The measurement system developed in this work is based on the open source hardware Arduino microcontroller, model Mega 2560 R3. Through this microcontroller we can read the data, display them on a screen and share them with other devices via Modbus.

The hardware consists of a printed circuit board with an ATmega2560 microcontroller. It includes 54 digital input/output ports, 16 analog inputs, 4 UARTs, a 16 MHz crystal oscillator, a power jack, an In Chip Serial Programmer (ICSP) header and a reset button. It also includes a USB connection port from where it is possible to power the board and establish communication with the computer.

In order to visualize the concentrations of the different gases that make up the biogas, an LCD graphic display will be used, which in fact acts as MI. It has been opted for a graphic LCD of 128x64 points with an ST7920 controller. For the connection of the LCD screen with Arduino, it has been decided to use the serial communication under the Serial Peripheral Interface (SPI) protocol.

The SPI bus is a communications standard, used mainly for the transfer of information between integrated circuits. It includes a line of clock, incoming data, outgoing data and a select chip pin, which connects or disconnects the operation of the device with which one wishes to communicate. In this way, this standard allows multiplexing the clock lines. Fig. 4 depicts the SPI bus scheme.

On the Arduino Mega 2560, the SPI port is located on pins 50 to 53.

The connection between the MI and the Arduino Mega are as shown below (Fig. 5).

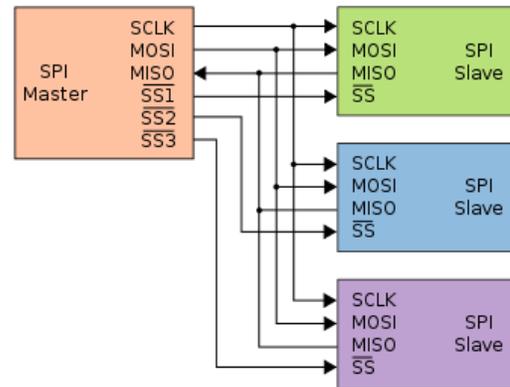


Figure 4. SPI bus scheme of connections.

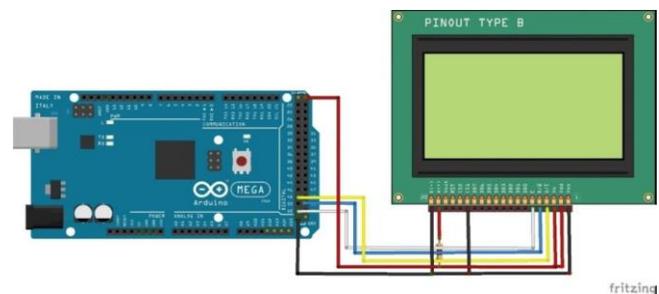


Figure 5. Connection between the MI and Arduino.

In order to use the Modbus TCP/IP protocol to access the sensor readings an Ethernet port must be added to the Arduino Mega. For this purpose, an Ethernet shield that fits perfectly with this version of Arduino will be used. Specifically, the Hanrun HR911105A Ethernet shield model was chosen, which uses the Wiznet W5100 Ethernet chip to allow Internet connection.

This Ethernet shield uses the SPI protocol to share the information with the Arduino board. Exactly, it uses the pins of the ICSP Port of the Arduino Mega for the connection. As can be seen in the Arduino Mega port scheme, the ICSP Port also allows connection with the SPI protocol. The physical assembly of this shield is very simple, it simply has to be inserted in the Mega Arduino and it will be connected and fixed.

Four already available commercial sensors of the manufacturer Figaro [34] are used to measure methane (model TGS3870), ammonia (model TGS2444), carbon monoxide (model TGS5042) and hydrogen (model TGS821). For further details about the features of these sensors and the auxiliary electronic circuitry, see reference [35].

Furthermore, apart from the Arduino board and the Ethernet shield and the LCD screen, to develop this prototype, a series of auxiliary electronic circuits are needed to operate the sensors. These auxiliary circuits are responsible for generating pulses, amplifying signals and sampling them.

B. Software implementation

For the implementation of the microcontroller script, known as sketch, the software consists of the Integrated Development Environment (IDE) based on the Processing environment and a programming language, a simplified version of the C++ language, as well as on the bootloader that is executed on the board. The Arduino IDE is a free software package that allows programming all Arduino boards. The microcontroller is programmed by a computer via the IDE and using serial communication via a RS-232 to TTL serial level converter. Moreover, open source libraries are used to configure the additional modules, in the present case, the Ethernet shield and the LCD screen.

The developed software consists on the programming of the data acquisition from the sensors, the conditioning and scaling of the measured signals and the presentation of the results in the LCD screen.

To control the graphic display, the U8g2 library has been chosen, a library that allows the use of a large number of monochrome displays with different controllers, including the one used in this work (ST 7920).

For the ModBus TCP/IP protocol, three libraries have been used; two of them are specific to Arduino IDE, such as SPI.h and Ethernet.h, which allow us to use the Ethernet protocol in this project. And the third one is Modbus Library for Arduino. This library implements the industrial communication protocol ModBus in Arduino, allowing the use of the Series physical layer (RS232 or RS485) or Ethernet (TCP/IP). Fig. 6 shows the beginning of the Arduino sketch where the libraries are included.

```
// Librerías utilizadas
////////////////////////////////////
#include <SPI.h>
#include <Ethernet.h>

#include <Modbus.h>
#include <ModbusIP.h>

#include <U8g2lib.h>

////////////////////////////////////
```

Figure 6. Libraries included in the sketch.

1) *Library U8g2*: Next, the operation of the U8g2 library will be briefly described. First, an object is defined to select the type of display, the size and type of buffer used. It is defined that the controller of the graphic LCD screen is the ST920: the screen resolution chosen has been 128 x 64 pixels. In addition, to remove load from the CPU, the hardware buffer that incorporates the controller has been used. The SPI port is used for the LCD communication and the Arduino pin 48 is defined as chip select line to enable the SPI slave. Once the library is configured, the *Starting_Screen ()* function will be executed within the program's *void setup ()*. This function is responsible for starting the screen and displaying an initial animation every

time the Arduino is started. Fig. 7 shows the corresponding code.

```
// FUNCIÓN DE ANIMACIÓN INICIAL. PANTALLA DE INICIO.

void Pantalla_Inicio () {

    u8g2.begin();
    u8g2.enableUTF8Print();

    u8g2.clearBuffer();
    u8g2.setFont(u8g2_font_10x20_tf);
    u8g2.drawFrame(0,0,128,63);
    u8g2.drawRect(0,60,128,3);
    u8g2.drawRect(125,0,3,63);
    u8g2.drawStr(12,17,"Analizador");
    u8g2.drawStr(50,35,"de");
    u8g2.drawStr(35,52,"Biogas");
    u8g2.sendBuffer();
    delay(1000);

}
```

Figure 7. Starting screen function of the U8g2 library.

2) *Modbus Library*: The operation of the Modbus library is concisely explained hereafter. First of all a ModbusIP object will be defined. This object will contain all the configuration and data registers. To make better use of the memory, the analog readings are scaled and transferred directly to the corresponding Modbus register. Once transferred to the Modbus registers, they will already be accessible from a remote master. Fig. 8 shows the corresponding code.

```
//// MODBUS
// Configuración de la conexión Ethernet
byte mac[] = { 0xDE, 0xAD, 0xBE, 0xEF, 0xFE, 0xED };

byte ip[] = { 192, 168, 1, 10 };

mb.config(mac, ip);

//TIPOS DE REGISTROS DISPONIBLES EN MB Y SUS CARACTERISTICAS
//mb.addCoil(x); // Reg "Coil" Digital out R/W
//mb.addHreg(x); // Reg "Holding" Analog out R/W
//mb.addIsts(x); // Reg "Input Status" Digital In R
//mb.addIreg(x); // Reg "Input Reg" Analog in R

for(int i=0;i<4;i++){
mb.addIreg(i);
}
```

Figure 8. Modbus library configuration.

IV. INITIAL OUTCOMES

The prototype is within a testing phase in order to evaluate its behavior and to detect improvements or corrections to carry out. In this sense, diverse experiments are being accomplished; some of them devoted to check the sensors' measurements whereas other trials are focused on the microcontroller performance. For instance, in the first group of experimentations, the repeatability of the

measurements is being studied. On the other hand, some delays are introduced into the Arduino sketch in order to establish the optimal sampling time.

A. Final assembly

In Fig. 9, the diagram of the final assembly is shown. The central element of this prototype is the Arduino Mega, with which all the peripherals communicate.

- LCD screen: Communicates with Arduino through the SPI bus.
- Ethernet Shield: Communicates with Arduino through the SPI bus.
- Gas sensors: Communicates with Arduino through analog signals between 0 and 5V.

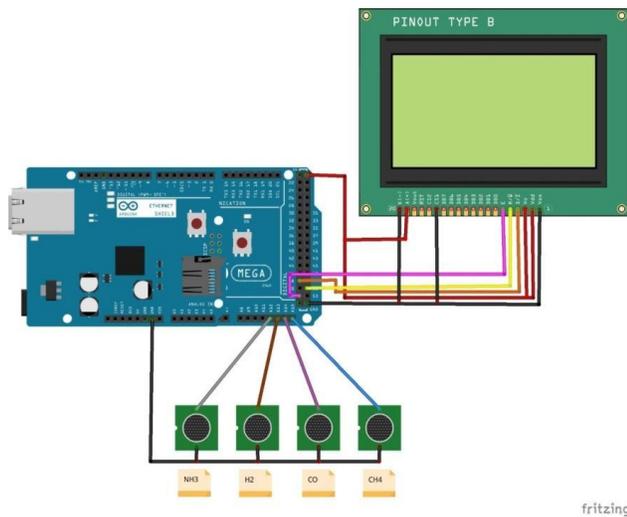


Figure 9. Diagram of the final assembly of the measurement system.

Fig. 10 contains a snapshot of the experimental measurement system once assembled in the laboratory approach.

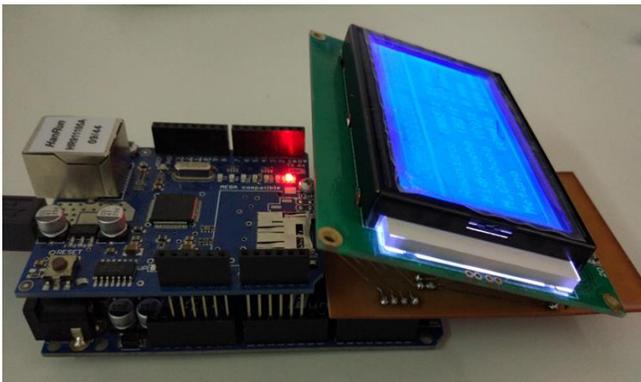


Figure 10. Final assembly of the experimental measurement system.

The physical assembling of the gas sensors is enabled by means of a custom-designed methacrylate box (Fig. 11), which hosts the sensors and makes the gas to circulate for being measured.



Figure 11. Custom-designed methacrylate box to host the gas sensors.

As a proof of concept, Fig. 12 shows a photograph of the prototype working, where can be appreciated that the MI provides real time data about the biogas composition.



Figure 12. Measurement system based on Arduino working.

B. Discussion and considerations

Once the prototype has been described, some considerations from the authors' perspective are commented below.

- Biogas sensing can be effectively solved using the open source microcontroller Arduino.
- The instrumentation required is already available but must be interfaced with the Arduino board with custom designed circuitry. The availability of shields devoted to connect sensors for gas measurement would reduce the time and effort of development.
- The wide range of connectivity interfaces provided by Arduino allows an easy development of networked communication.
- As a consequence of the available information about Arduino-related projects, the effort necessary to become familiar with the microcontroller resources has been relatively short. In fact, the time devoted to

configure the Arduino board has been shorter than that required to configure the sensors.

- Further improvements of Arduino libraries and/or shields will enrich the proposed system.
- Some minor drawbacks have been found during the development such as a weak physical connection between the GLCD and the main board, as well as inaccuracy of the sample time achieved by the microcontroller.
- The prototype must be assessed under real operation conditions in order to check its robustness.
- Reliability and repeatability of the measurements must be studied before an intensive utilization of the prototype.

V. CONCLUSIONS

A measurement system based on the Arduino microcontroller devoted to biogas sensing has been presented. A set of sensors are responsible of measuring four gases, namely hydrogen, methane, ammonia and carbon monoxide. On the other hand, the Arduino board performs the signal processing of the acquired data and establishes the percentage of each gas, displaying such information by means of a small screen, which acts as MI.

This system is still under development and validation; nonetheless, initial outcomes at laboratory scale demonstrate the suitability of Arduino microcontroller for biogas sensing.

Arduino acts as the core of the developed measurement system, providing important advantages like low cost, open source hardware, great amount of information available in the Internet and easy configuration, just to name a few. As it is evident, this device also presents some drawbacks, being the main one the inaccuracy of the sampling time.

Further works are focused on integrating the proposed approach with a monitoring and supervisory system in order to enable real time visualization of the biogas composition. In addition, the networked operation to provide cloud-enabled measurements storage is also being considered.

ACKNOWLEDGMENT

Authors are grateful to the community that supports Arduino-based developments under open source philosophy.

REFERENCES

- [1] A.J. Calderón and I. González, "Design and Implementation of a Low Cost System to Determine the Composition of Biogas," Proc. Eighth International Conference on Sensor Device Technologies and Applications (SENSORDEVICES 2017), IARIA, Sep. 2017, pp. 76-77, ISBN: 978-1-61208-581-4.
- [2] D. Deublein and A. Steinhäuser, *Biogas from Waste and Renewable Resources: An Introduction*, Wiley-VCH, 2011.
- [3] I. González, A.J. Calderón, J.M. Andújar, "Novel Remote Monitoring Platform for RES-Hydrogen based Smart Microgrid," Energy Conversion and Management, vol. 148, pp. 489-505, 2017, doi: 10.1016/j.enconman.2017.06.031.
- [4] R. Fisher, L. Ledwaba, G. Hancke, C. Kruger, "Open Hardware: A Role to Play in Wireless Sensor Networks?," Sensors, vol. 15, pp. 6818-6844, 2015, doi: 10.3390/s150306818.
- [5] B. Martínez, X. Vilajosana, I.H. Kim, J. Zhou, P. Tuset-Peiró, A. Xhafa, "I3Mote: An Open Development Platform for the Intelligent Industrial Internet," Sensors, vol. 17, pp. 986, 2017, doi: 10.3390/s17050986.
- [6] A. Garrigós, D. Marroquí, J.M. Blanes, R. Gutiérrez, I. Blanquer, M. Cantó, "Designing Arduino electronic shields: experiences from secondary and university courses," Proc. IEEE Global Engineering Education Conference (EDUCON), April 2017, doi: 10.1109/EDUCON.2017.7942960.
- [7] Arduino homepage. Available from: <https://www.arduino.cc/> (last accessed February 2018)
- [8] A.J. Calderón, I. González, M. Calderón, F. Segura, J.M. Andújar, "A New, Scalable and Low Cost Multi-Channel Monitoring System for Polymer Electrolyte Fuel Cells," Sensors, vol. 16, pp. 349, 2016, doi: 10.3390/s16030349.
- [9] A. Cela, J.J. Yebes, R. Arroyo, L.M. Bergasa, R. Barea, E. López, "Complete low-cost implementation of a teleoperated control system for a humanoid robot," Sensors, vol. 13(2), pp. 1385-1401, 2013, doi:10.3390/s130201385.
- [10] C. Robles, J. Callejas, A. Polo, "A. Low-Cost Fuzzy Logic Control for Greenhouse Environments with Web Monitoring," Electronics, vol. 6, pp. 71, 2017, doi: 10.3390/electronics6040071.
- [11] L.B. Souza, P.P. Soares, R.V. Duarte, M. Mendonça, E.I. Papageorgiou, "Dynamic Fuzzy Cognitive Maps Embedded and Intelligent Controllers Applied in Industrial Mixer Process," International Journal On Advances in Systems and Measurements, vol. 10, numbers 3 and 4, pp. 222-233, 2017.
- [12] M.A. Prada, P. Reguera, S. Alonso, A. Morán, J.J. Fuertes, M. Domínguez, "Communication with resource-constrained devices through MQTT for control education," IFAC-PapersOnLine, vol. 49(6), pp. 150-155, 2016, doi: 10.1016/j.ifacol.2016.07.169.
- [13] I. Tejado, J. Serrano, E. Pérez, D. Torres, B. Vinagre, "Low-cost Hardware-in-the-loop Testbed of a Mobile Robot to Support Learning in Automatic Control and Robotics," IFAC-PapersOnLine, vol. 49(6), pp. 242-247, 2016, doi: 10.1016/j.ifacol.2016.07.184.
- [14] A. Mejías, M. Reyes, M.A. Márquez, A.J. Calderón, I. González, J.M. Andújar, "Easy handling of sensors and actuators over TCP/IP Networks by Open Source Hardware/Software," Sensors, vol. 17, pp. 94, 2017, doi: 10.3390/s17010094.
- [15] J. Chacón, J. Saenz, L. de la Torre, J.M. Diaz, F. Esquembre, "Design of a Low-Cost Air Levitation System for Teaching Control Engineering," Sensors, vol. 17, pp. 2321, 2017, doi: 10.3390/s17102321.
- [16] R. Pereira, J. Figueiredo, R. Melicio, V.M.F. Mendes, J. Martins, J.C. Quadrado, "Consumer energy management system with integration of smart meters," Energy Reports, vol. 1, pp. 22-29, 2015, doi:10.1016/j.egy.2014.10.001.
- [17] M.V. García, E. Irisarri, F. Pérez, E. Estévez, M. Marcos, "OPC-UA Communications Integration using a CPPS architecture," Proc. Ecuador Technical Chapters Meeting (ETCM), Oct. 2016, doi: 10.1109/ETCM.2016.7750838.
- [18] M. Müller, E. Wings, L. Bergmann, "Developing open source cyber-physical systems for service-oriented architectures using OPC UA," Proc. IEEE 15th International Conference on Industrial Informatics (INDIN), Jul. 2017, doi: 10.1109/INDIN.2017.8104751.
- [19] M. Kajan, J. Sovcik, F. Duchon, L. Mrafko, M. Florek, P. Beño, "Sensoric subsystem of automated guided vehicle: TCP communication between SIMATIC S7 PLC and Arduino," Proc. 23rd International Conference on Robotics in Alpe-

- Adria-Danube Region (RAAD), Sep. 2014, doi: 10.1109/RAAD.2014.7002258.
- [20] H. Rahadian, W. Nugroho, D.Nurual, "Siemens CPU1215C Input Expansion and Remote Monitoring with Arduino Bridge," Proc. 6th International Annual Engineering Seminar (InAES), Aug. 2016, doi: 10.1109/INAES.2016.7821926.
- [21] U. B. Kayande, P. Jaspreetkaur, R. M. Kulkarni, S. B. Mahajan, P. Sanjeevikumar, "Digitally Controlled Hybrid Liquid Level Detection System Using Programmable Logic Controller and Microcontroller," Lecture Notes in Electrical Engineering, vol. 442, pp. 365-374, 2018 doi: 10.1007/978-981-10-4762-6_35.
- [22] B.D. Lampson, Y.J. Han, A. Khalilian, J.K. Greene, D.C. Degenhardt, J.O. Hallstrom, "Development of a portable electronic nose for detection of pests and plant damage," Computers and Electronics in Agriculture, vol. 108, pp. 87-94, 2014, doi: /10.1016/j.compag.2014.07.002.
- [23] D. Piromalis, K. Arvanitis, "SensoTube: A scalable hardware design architectures for wireless sensors and actuators network nodes in the agricultural domain," Sensors, vol. 16, pp. 1227, 2016, doi: 10.3390/s16081227.
- [24] N.C. Batista, R. Melício, J.C.O. Matias, J.P.S. Catalão, "Photovoltaic and wind energy systems monitoring and building/home energy management using ZigBee devices within a smart grid," Energy, vol. 49, pp. 306–315, 2013, doi: /10.1016/j.energy.2012.11.002.
- [25] P. Papageorgas, D. Piromalis, K. Antonakoglou, G. Vokas, D. Tseles, K.G. Arvanitis, "Smart Solar Panels: In-situ monitoring of photovoltaic panels based on wired and wireless sensor networks," Energy Procedia, vol. 36, pp. 535-545, 2013.
- [26] H.E. Gad and H.E. Gad, "Development of a new temperature data acquisition system for solar energy applications," Renewable Energy, vol. 74, pp. 337-343, 2015, doi: 10.1016/j.renene.2014.08.006.
- [27] M.M. Rahman, J. Selvaraj, N.A. Rahim, M. Hasanuzzaman, "Global modern monitoring systems for PV based power generation: A review," Renewable and Sustainable Energy Reviews, vol. 82(3), pp. 4142-4158, 2018, doi: 10.1016/j.rser.2017.10.111.
- [28] J. Carland, M. Umeda, T. Wilkey, A. Oberbeck, J. Cumming, N. Parks, "Meteorological Wireless Sensor Networks," Sensors & Transducers, vol. 160, pp. 118-124, 2013.
- [29] S. Ferdoush and X. Li, "Wireless Sensor Network System Design Using Raspberry Pi and Arduino for Environmental Monitoring Applications," Procedia Computer Science, vol. 34, pp. 103–110, 2014.
- [30] S. Trilles, A. Calia, O. Belmonte, J. Torres-Sospedra, R. Montoliu, J. Huerta, "Deployment of an open sensorized platform in a smart city context," Future Generation Computer Systems vol. 76, pp. 221-233, 2017, doi: 10.1016/j.future.2016.11.005.
- [31] P. Arroyo, J. Lozano, J.I. Suarez, J.L. Herrero, P. Carmona, "Wireless Sensor Network for Air Quality Monitoring and Control," Chemical Engineering Transactions, vol. 54, pp. 217-222, 2016.
- [32] A. Arenal, J. Pascual, R. González, S. Ríos, "Platform for controlling and getting data from network connected drones in indoor environments," Future Generation Computer Systems, 2018, In press, doi: 10.1016/j.future.2018.01.011.
- [33] F. Segura, V. Bartolucci, J.M. Andújar, "Hardware/Software Data Acquisition System for Real Time Cell Temperature Monitoring in Air-Cooled Polymer Electrolyte Fuel Cells," Sensors, vol. 17, pp. 1600, 2017, doi: 10.3390/s17071600.
- [34] FIGARO ENGINEERING INC. homepage. Available from: <http://www.figaro.co.jp/en/product/> (last accessed April 2018)
- [35] A.J. Calderón and I. González, "Biogas analyzer based on Open Source Hardware: Design and Prototype Implementation," Sensors & Transducers, vol. 220, pp. 31-36, 2017.

Low Power Optimized and DPA Resistant D-FF for Versatile Mobile Applications

Karol Niewiadomski, Dietmar Tutsch
 University of Wuppertal
 Chair of Automation and Computer Science
 Wuppertal, Germany
 Email: {niewiadomski, tutsch}@uni-wuppertal.de

Abstract—Starting from early, simple logic gates, the development of integrated electronics made impressive proceedings in terms of performance and complexity. These chips can be found everywhere in daily life, serving the purpose of processing tasks, which are mostly too complicated, dangerous or tiring to human's nature. Whilst the use of integrated circuits was limited to classic applications, e.g., personal computers, servers, mainframes, etc., the application scope was continuously enlarged over the years. Hence, microprocessors began to add more and more computational power to various mobile devices, e.g., cell phones, tablets and even vehicles. Nowadays, smartphones provide more processing resources than bulky mainframes used for ballistic calculations ever did. It is thinkable that smartphones and tablets will replace the classic personal computers in many households, due to the mobility, versatility and simplicity they offer. Another example for fast digitalization are vehicles. Next generation cars will offer a growing number of automated driving assistance systems, which shall add safety and comfort to daily traffic situations. Further steps towards vehicle to vehicle and vehicle to infrastructure communications will produce tremendous amounts of data. As a consequence, more processing capabilities will be needed over time and therefore challenging the lifetime of batteries. Architectural improvements towards battery lifetime extension are an inevitable step, however, power sensitive adaptations must be done at a deeper hierarchical level. Since each data processing logic heavily depends on registers implemented by data flip-flops, this paper presents a newly designed charge recycling data flip-flop. Major focus during research and development was put on low power design aspects as well as on security-related enhancements to counter differential power analysis. This new design is compared to a selection of various, already existing implementations.

Keywords—FPGA; D-FF; charge recycling; low-power; differential power analysis.

I. INTRODUCTION

Mobile applications like notebooks, smartphones, tablets and wearables have changed the usage behavior over the last years. The access to information shall be available everywhere and completely independent from classic computers. This trend can be clearly seen in the current digitalization of vehicles, providing more and more features like driving assistance systems and interfaces for the connection of smartphones for displaying installed apps on the embedded infotainment system [1]. A modern, upper-class vehicle contains more than 70 electronic control units (ECUs) to provide all features desired by consumers these days [2]. Such applications rely on the provision of sufficient processing power, which in turn requires adequate energy resources. Both, handheld computation units and vehicles have only limited battery capacities, therefore, a necessity for power optimized integrated electronics is given.

One approach to overcome these challenges are FPGAs. These integrated circuits play a major role for the realization of adaptive and efficient systems, offering vast reconfiguration abilities [3] [4]. Reconfigurability goes back on arrays of memory cells like static random access memory (SRAM). In order to optimize an FPGA in terms of energy efficiency, these memory cells have to be extended with power reduction measures [5] [6]. In addition to that, each FPGA works with flip-flops, which have an influence on the overall speed of the design since they are driven by the system clock. Furthermore, approximately 30% - 70% of the total power in a clocked design is dissipated by the clocking network, which is absolutely crucial for the operation of these circuits [7]. In consequence, by carefully re-designing these commonly used D-FFs, energy consumption can be decreased by applying static and dynamic power reduction measures. Power constraints are one of the most important challenges in modern circuit design. In addition, cyber security has become a frequently discussed topic in recent years, due to many incidents and a rising awareness for data protection. Side-channel attacks, which are based on differential power analysis, illustrate a possibility how to reveal confidential data without physical access to critical devices [8]. Thus, dedicated circuit modifications at circuit level shall be used for catching potential threats.

In this paper, we investigate selected D-FF cell designs on their low power characteristics, which can not be neglected in battery-powered systems. In Section II, we give an overview about related work, which includes a selection of existing D-FF designs on their assets and drawbacks, as well as key aspects of dependencies between performance and power consumption. In Section III, we present our charge recycling (CR) D-FF and explain the implemented circuit improvement methods for static and dynamic power reduction. In Section IV, we discuss simulation results of the D-FF and analyze the benefits of power reduction measures based on these simulations. In Section V, all previous discussions are summarized and concluded.

II. RELATED WORK

In general, we can distinguish between two different types of storage elements used in registers of, e.g., processors: latches and flip-flops (FF). Both designs inhibit their pros and cons and are typically designed to serve different purposes, which shall be illustrated in Figure 1. The inputs for both implementations are a clock signal T and an input signal

D , which can be a sequence of pre-defined voltage levels based on a randomly generated sequence of input data. A latching circuit shows full transparency once T is put to *HIGH*: regardless what kind of logic value is applied on D , the output of a latch follows every change on its input node immediately. On the other hand, once T is driven to *LOW*, this kind of circuit latches or stores the latest input applied on D before the clock signal is changed from 1 to 0. Depending on the respective applications, this special feature called transparency might be a desirable behavior or not. To overcome this problem and to have a real alternative to latches, FF circuits were invented, which are sensitive to the rising or falling edge of the clock signal. In special cases, even both edges can be used to evaluate the applied data.

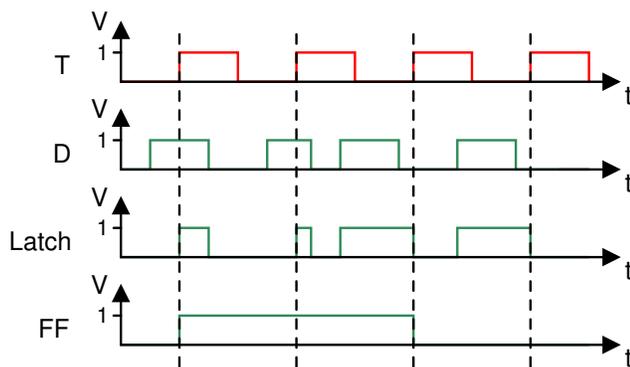


Fig. 1. Basic working principle of latches and D-FFs

Figure 1 shows the function of a positive-edge driven FF. As the edge of T rises from 0 to 1, the FF samples the data D and stores it until the next rising edge of T , regardless of all changes on D . Hence, the transparency effect of latching circuits is avoided. Furthermore, an additional evaluation of D could be implemented to sampling the input signal even during the falling edge of T , which leads to faster operation. However, this comes along with some modifications and should be rather decided case by case. Certain design offer the possibility to be configured either to work as latch or as FF, but since the transparency effect is of no benefit in many cases, this paper focuses on the investigation of a low-power FF.

D-FFs are the working horse in different applications, like storage registers, counters, frequency dividers, etc. FPGAs resort on these circuits in each slice, which is a basic computational element, shown in Figure 2.

Each slice contains one D-FF for storage of computed values prior to forwarding them to the next configurable logic block (CLB). Since even a low-cost FPGA, e.g., Xilinx Spartan 3A, contains up to 8320 CLBs [9], one can see the strong impact on area and energy consumption of these clocked devices. The relation between consumed power and the supply voltage, load capacitance and system clock can be seen in (1):

$$P = \alpha CV^2 f_{Clk} \quad (1)$$

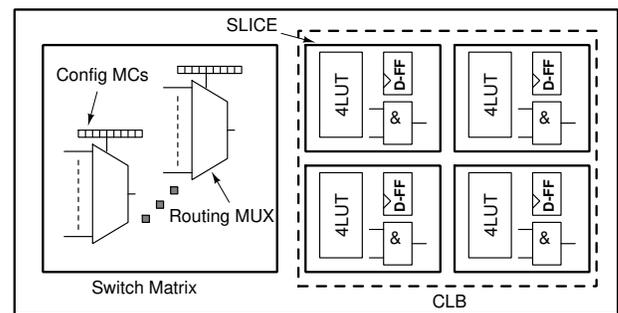


Fig. 2. Simplified SLICE structure of an FPGA

The activity factor α represents the cadence of write requests. A reduction of α can be achieved by special memory cell designs [10] or alternatively with auxiliary comparator circuitry. Another efficient approach is reducing the operating voltage. This can be achieved by techniques like dynamic voltage scaling (DVS), which was evaluated in various publications [11]. Power gating is certainly the strongest way to achieve a measurable reduction of energy consumption. However, this can be only applied, if there is no focus on data retention. A further possibility for raising the energy efficiency is lowering the clock frequency f_{Clk} . Circuitry, which is not timing critical can be clocked down to a minimum speed, which ensures a reliable operation of the system. If certain circuit parts can be completely stopped while retaining stored logic values, full clock gating can be a feasible solution to save power [12]. Both methods can be combined on a coarse-grain or fine-grain level.

These techniques are only an extract of a set consisting of different methods on how to handle the challenges of demanding functions. A majority of these solutions require additional circuitry to be added and implemented at a higher architectural level. Our approach goes one step further and is based on direct circuit level improvements to a D-FF by reasonable selection of a suitable D-FF cell design and substantial modifications of the internal cell circuitry to achieve better efficiency. The improvements achieved on that level are essential for important energy dissipation suppression and are an inevitable step for optimization to be combined with architectural amendments.

Different concepts have been introduced in the recent years. In general, we can distinguish between latches and flip-flops. Whilst latches are level-sensitive designs, flip-flops are edge-sensitive. Latches are transparent and therefore not suitable for timing-critical applications due to possible glitches in the signal path. For avoiding glitches and in consequence timing problems in complex designs, many flip-flop designs implicate the principle of cascading master-slave D-FFs. This standard design is shown in Figure 3.

Both, master and slave unit, consist of a feedback loop of inverters and transmission gates. Once Clk is set to *HIGH*, the input data provided by D is latched in the master circuit. At this point, the transmission gate connecting master and slave circuit, is in cut-off mode and therefore avoiding any

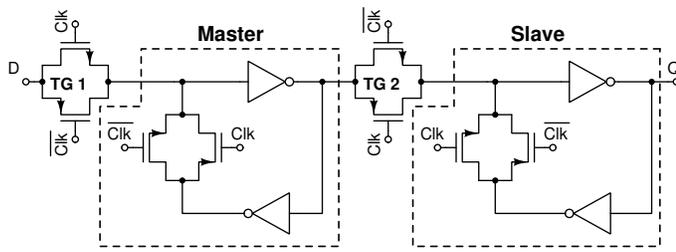


Fig. 3. Master-slave arrangement

glitches, e.g., direct throughput of D to Q . When Clk is set to LOW , the stored data at the output of the master circuit is latched by the subsequent slave unit and provided at the output node Q . Any changes of D will not influence the logic value stored at Q due to the fact that both transistors of TG 1 are in cut-off mode. This legacy design was the starting point for numerous variations in the past. All simulations have been performed with Cadence tools and a $90nm$ technology provided by TSMC at an ambient temperature of $27^{\circ}C$. The clock frequency was set to $250MHz$.

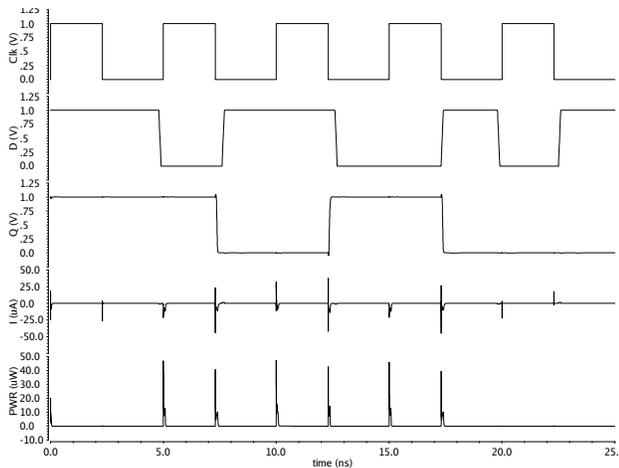


Fig. 4. Simulation results of SET D-FF

1) *SET D-FF*: A simplified implementation is shown in Figure 5. Whilst the reference design of a D-FF uses 16 transistors in total, this design consists of 10 transistors only, leading to a higher chip density and reduced manufacturing costs [13].

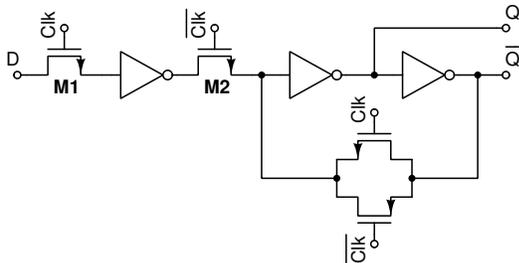


Fig. 5. Single Edge Triggered D-FF

Instead of 4 TGs, this design works with 1 TG and achieves the same function by replacing the remaining TGs by nMOS transistors. This reduction of transistors comes along with cutting down the number of slower and larger pMOS transistors. Furthermore, this implementation provides the generation of both Q and \bar{Q} . The functionality of the SET D-FF is similar to the reference design: glitching is avoided by complementary control of both pass-transistors $M1$ and $M2$. Latching and generation of the output values is done in the feedback loop after the activation of $M2$. Analog to the previous standard design, this concept relies on the preparation of complementary Clk signals, which requires additional circuitry for signal generation. The simulation results are illustrated in Figure 6. The slew rate of Q of during a $HIGH \rightarrow LOW$ switching event is noticeable weaker than of its \bar{Q} counterpart. This goes back to the additional inverter, which is placed right after the node where Q is generated in the signal path. This inverter is used to achieve a higher slew rate of \bar{Q} but also adds a slight delay. To improve the slew rate of Q an appropriate adaption of the signal chain's second inverter transistor parameters should be done.

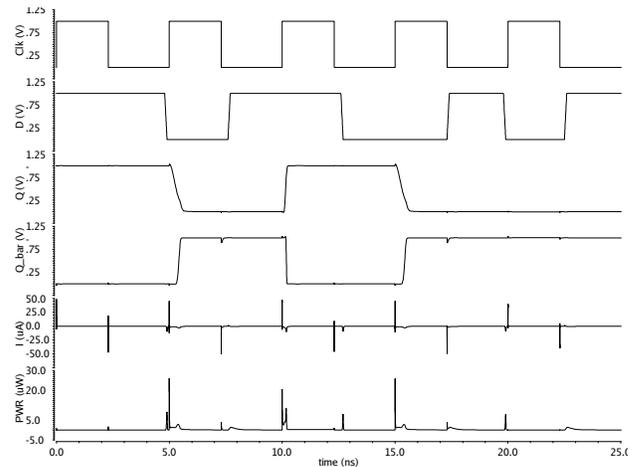


Fig. 6. Simulation results of SET D-FF

2) *Low-power D-FF*: Another variation, which displays an attempt on how to optimize a D-FF with respect to power consumption, is shown in Figure 7. The key aspect of this design is to eliminate short-circuit power dissipation from the feedback path [14] due to the tri-state inverter. Although keeping the same number of transistor like in the reference design, considerable power savings can be achieved. This will be discussed in the last section of this paper.

In direct comparison to the SET D-FF, Figure 8 depicts a better slew rate of the output signal Q , regardless of considering a $HIGH \rightarrow LOW$ or $LOW \rightarrow HIGH$ transition. However, this design does not support provision of complementary outputs, which would come along with further modifications.

3) *PPI D-FF*: In order to get a better performance of a conventional D-FF, the Push-Pull-Isolation (PPI) D-FF was presented in [14]. The main advantage of this implementation is the reduced clock-to-output delay from two gates in the

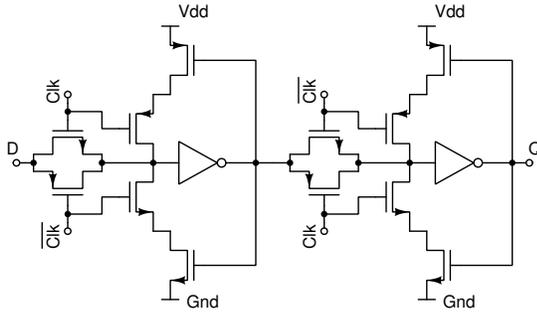


Fig. 7. Low-power modification of D-FF

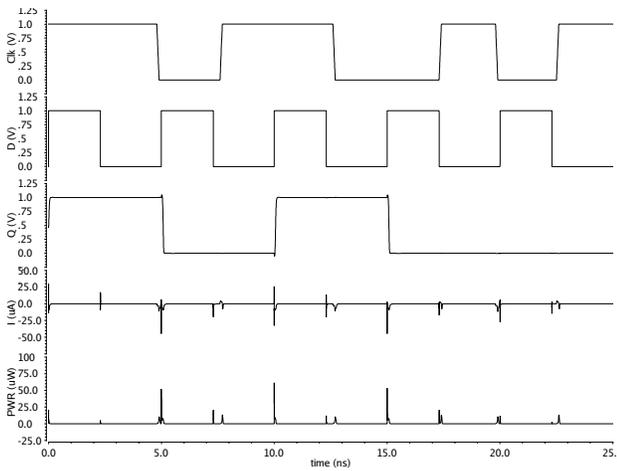


Fig. 8. Simulation results of Low-power D-FF

reference design to one gate in the PPI D-FF, which is shown in Figure 9.

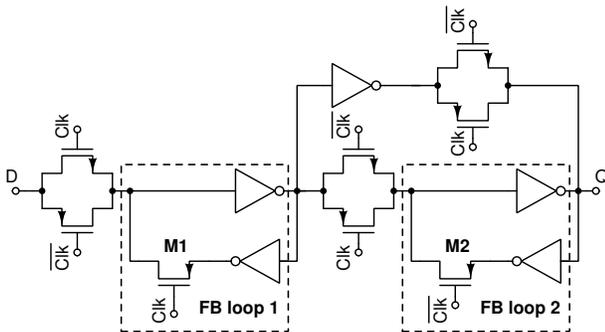


Fig. 9. Push-Pull-Isolation D-FF

The insertion of an inverter and a TG between the output nodes of master and slave latches provides a push-pull effect at the slave latch. In consequence, the input and output of the inverter in the slave unit will be driven to opposite logic values during operation. This design is approximately 31% faster than the reference D-FF, but has a power overhead of 22%. To counter the increased power consumption 2 pMOS transistors, *M1* and *M2*, are added to the feedback loops in the master and slave latches. In direct comparison with the conventional D-FF, the PPI D-FF improves speed by 56% at an expense of 6% of additional power dissipation. The respective simulation

results are shown in Figure 10.

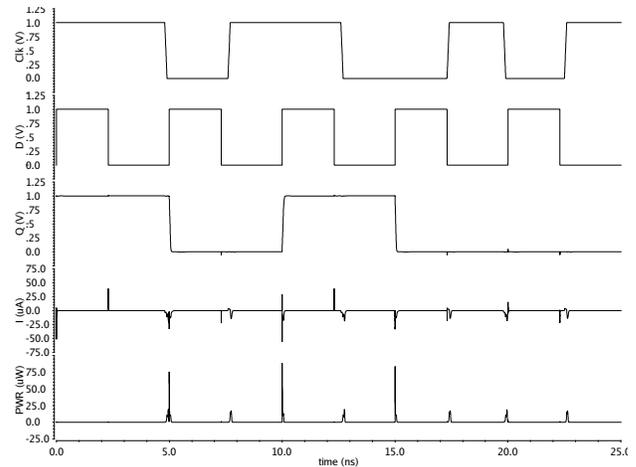


Fig. 10. Simulation results of PPI D-FF

For achieving comparable results, all designs have been simulated with the same test circuit and same stimuli inputs. The related simulation environment is shown in Figure 11. The input signals *D*, *Clk* and \overline{Clk} are provided by the driver circuitry. Since the signal transition through a simple inverter adds some delay between both *Clk* signals, additional circuitry for synchronizing these signals must be added. For the sake of simplicity, this is not shown in Figure 11. The load consists of 2 capacitors of 200fF, emulating parasitic capacitances of the metal layers and 2 additional inverters at the output nodes. The design under test (DUT) is powered by an independent voltage source to enable a precise comparison of the D-FF designs in scope of this paper. For the low power and PPI D-FF, which are not supporting the generation of \overline{Q} , the test circuit has been appropriately adapted.

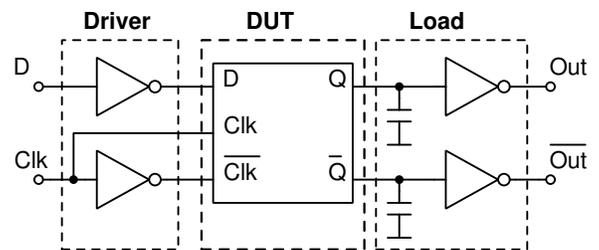


Fig. 11. Test circuit

For all introduced cell designs in this paper, the average power consumption, the maximum and minimum power consumption during simulation time were traced and summarized in Table I. These results show that the reference D-FF dissipates the highest average power consumption by 1186nW, due to lack of power savings measures. The maximum power dissipation confirms this result by revealing a higher consumption by the factor of approximately 4 in direct comparison with the optimized low-power D-FF. However, this result was expected and highlights the improvements of previously introduced designs.

TABLE I. SIMULATION RESULTS (PWR)

D-FF Type	Average Power nW	Max. Power μ W	Min. Power fW
Reference	1186	233.3	51.47
SETD	280.3	26.21	22.39
Low-power	272.7	61.55	19.92
PPI	435.4	88.71	28.01

On the other hand, similar results are reflected by measuring the leakage current of each design, shown in Table II. The reference D-FF exhibits the highest average leakage current I_{leak} by $1262nA$, which is approximately fivefold higher than average I_{leak} of the low-power D-FF. Analog to the average leakage current, the maximum leakage current is also allocated to the reference design and points out that all power-optimized variations perform better in terms of energy efficiency.

TABLE II. SIMULATION RESULTS I_{Leak}

D-FF Type	Avg. Current nA	Max. Current μ W	Min. Current μ W
Reference	1262	336.3	346
SETD	265.7	48.94	50.41
Low-power	235.1	28.83	45.9
PPI	403.7	39.4	56.35

The respective simulation results are shown in Figure 12, which illustrates the input signal D , the clock signal Clk and the respective power dissipation output profiles for the presented input sequence with an alternating $0 \rightarrow 1 \rightarrow 0 \rightarrow 1$ sequence.

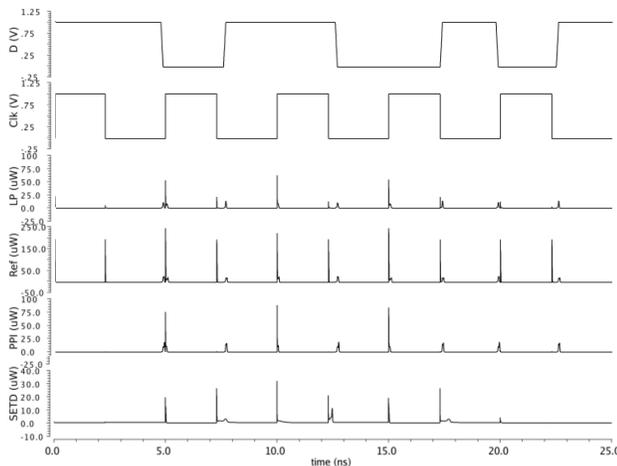


Fig. 12. Comparison Results

All designs exhibit strongly varying power consumption for each transition on the input nodes during the rising edge of the Clk signal, which comes along with an exploitable vulnerability for side-channel attacks. Glitches can be identified during the falling edge of Clk , which indicates weaknesses in the latching mechanism of master and slave latch, therefore revealing undesired transparency. None of the previously presented designs is optimized in terms of static leakage current suppression or energy recovery during runtime, which will be key aspects of our presented design in the next section.

III. CR D-FF

Based on the analysis of drawbacks of existing D-FF designs, we present a new approach of a low-power, energy-efficient and glitch-free D-FF, which is suitable for security-relevant applications with limited energy resources. Referring to the standard design shown in Figure 3, our intention was to redesign a new flip-flop cell from scratch. Without any direct relation to the D-FFs presented in the previous section, we present our charge recycling (CR) D-FF, which is illustrated in Figure 13. The transistors $M3$ and $M6$ are turned on during a LOW phase of the CLK signal and therefore charging both output nodes to V_{dd} . This procedure can be seen as a drawback in this design since it puts a strain on the power supply / battery of a mobile device. Hence, further optimization shall be done for achievement of an effective relief of available energy resources and therefore leading to a series of necessary adaptations.

This newly implemented design features a series of dedicated power savings mechanisms, which will be discussed in the following sections.

A. Charge recycling

Storing and processing logic values in flip-flops, registers, memories leads to charging and discharging of parasitic capacitances, which are an essential part of each integrated circuit. The development of a Sense Amplifier Based Logic (SABL) D-FF was an intermediate step towards the development of the CR D-FF. An implementation of a simple SABL inverter is shown in Figure 14 and the respective simulation results in Figure 15. The simulation curves show the correct functional behavior of this inverter and its special characteristic during operation: alternating precharge ($CLK LOW$) and evaluation ($CLK HIGH$) phases. One essential benefit of this design is the almost equal power dissipation during both phases, which adds essential value to countering DPA attacks. This can be seen by evaluation of the current spikes in Figure 15.

Since the CR D-FF features dynamic logic, periodic charge & discharge cycles are an integral part of the intended function and require special attention during the design. Similar to the introduced SABL inverter this design works with 2 alternating phases during runtime: precharge & evaluate, which are both triggered by the Clk signal. Whilst Clk turns to LOW , $M5$ is turned on and in consequence also switching on the pMOS transistors $M3$ & $M6$. Illustrating a critical point with respect to power savings within an integrated circuit, the precharge phase is the more deciding one. Due to the fact that these transistors are therefore in a conducting state, the capacitances at the output nodes Out & \bar{Out} are shortened. Hence, not discharged electrons at one of the complementary output nodes are used for charging the previously discharged output node. This effect is used for equilibrating electron charges and thus relieving the battery due to the fact that less energy is needed. This is a strong method for achieving a better performance in terms of dissipation reduction during dynamic behavior.

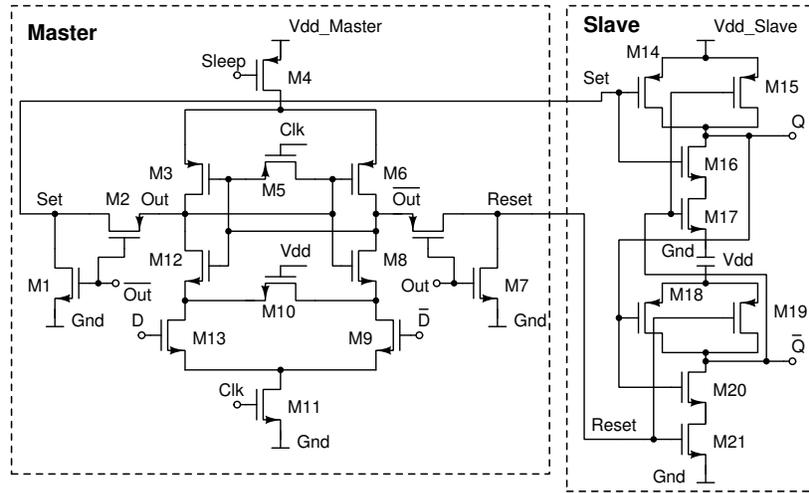


Fig. 13. CR D-FF

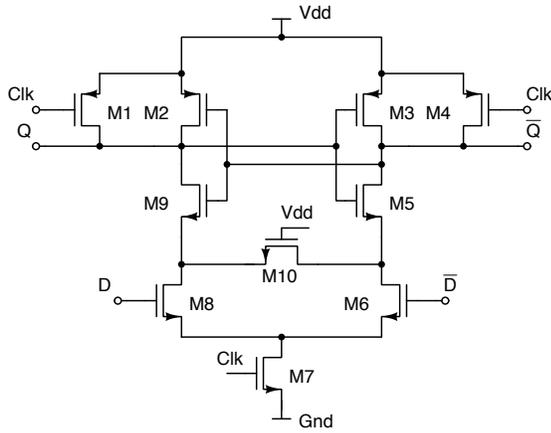


Fig. 14. SABL Inverter

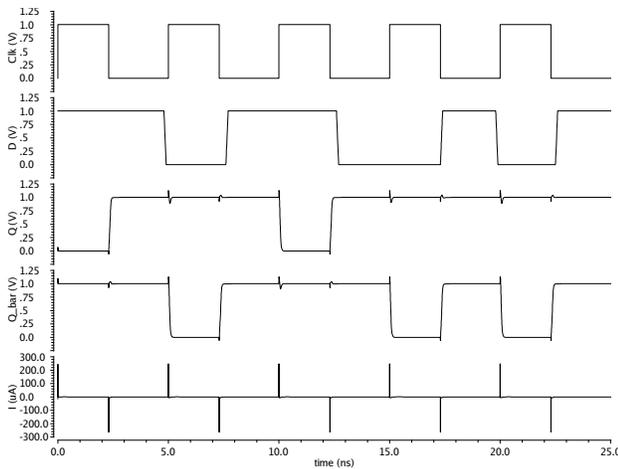


Fig. 15. SABL Inverter Simulation Results

After *Clk* applies a logic *HIGH* at the gate of *M4*, this transistor is turned off whereas *M11* is turned on and subsequently starting the evaluation phase in terms of sensing

the difference between the complementary inputs *D* & \bar{D} . One of the various benefits of sense amplifier based logic is that even a small Δ voltage between both input signals will be sensed and evaluated, providing a higher speed of the D-FF.

B. Dual Threshold CMOS

Leakage currents I_{leak} during standby contribute to a significant amount of total dissipation loss. By adding dedicated countermeasures, appreciable power savings can be achieved without investing much effort for realization. This can be done by the usage of transistors with a high threshold voltage V_{th} . Transistors with a high V_{th} require a proportional higher V_{GS} voltage at their gate nodes in order to be turned on, which implies a mitigation of leakage currents. This method can be combined by applying a negative V_{GS} for leading transistors into a deep turn-off status and therefore supporting suppression of leakage currents. This technique should be only applied carefully on circuit parts, which are not timing-critical since higher threshold voltages usually equal in slower signal transition. All transistors in our design are high V_{th} transistors for the sake of strongest suppression of I_{leak} .

C. Multi-oxide technology

Closely related to the previous section, static power dissipation can be further decreased by improving the tunneling-barrier for electrons. Undesired tunneling of electrons through the gate to bulk leads to current flows, which shall be eliminated. The relation between I_{leak} and the tunneling-barrier is shown in (2):

$$I_{leak} \propto A \left(\frac{V_{ox}}{T_{ox}} \right)^2 \tag{2}$$

Increasing the tunneling-barrier can be realized by increasing the gate oxide thickness T_{ox} . A higher oxide thickness leads immediately to a reduction of the tunneling current density I_{leak} , following the goal to extend battery lifetime

of mobile devices even in standby mode. The drawback of this technique is similar to the previous one: penalty of the circuit speed may occur if not applied carefully. Based on this reason, we decided to use high T_{ox} transistors for $M4$, $M5$ and $M11$. All of these transistors are not timing-critical, since $M4$ is used to activate a dedicated sleep mode and $M5$ for balancing the outputs. All of these functions are not slowing the circuit speed.

D. Clk- and power-gating

For further reduction of dynamic power dissipation, cutting off the Clk signal leads to transfer the circuit to a hold state, while maintaining the stored data inside the latches. Circuitry, which is not executing different operations over runtime, can be kept in a *WAIT* state, ready to continue calculation whenever the Clk signal is set to *HIGH* again. In the proposed design, $M5$ & $M11$ are used for stopping the D-FF from operating, but still keeping the correct data at the outputs of the cross-coupled inverters. Of course, additional circuitry driving and distributing the Clk signal over a whole design is an indispensable requirement. This can be provided by digital clock managers (DCMs), which are not covered by this paper.

In case that data storage is not necessary, gating of the supply voltage is an effective method how to save power in unused parts of a circuit. Power gating can be applied on different hierarchical levels. Our decision was to follow a fine-grain approach, leading to equipping the proposed D-FF with a power gating transistor $M4$. If the *SLEEP* signal turns from 0 to 1, $M4$ is off and therefore disconnecting the D-FF from V_{dd} . If this technique is applied in accordance with clock gating, total rail-to-rail-decoupling (V_{dd} & Gnd) can be realized.

E. Stacking

Transistor stacking is a further, strong technique for sub-threshold current reduction. Stacking transistors means to increase to source voltage V_S while keeping the gate voltage V_G at the same level. At a certain point of time, V_{GS} becomes negative, which leads the transistor into super cut-off mode and turns it deeply off. The more transistors are stacked in series, the better leakage current reduction will be. However, the most significant results can be achieved by adding a second transistor in series, because the effect of subthreshold current reduction becomes diminished with a rising number of transistors. Our proposed D-FF features stacking as a design principle, e.g., in the pull-down-networks of the slave latch, realized by $M16$ $M17$ and $M20$ & $M21$.

IV. SIMULATION RESULTS

As a starting point for further considerations and a better comparison, a CR inverter was implemented, shown in Figure 16. The total number of used transistors for this inverter's design is 9 and therefore, 1 transistor less than compared to the SABL implementation. Figure 17 shows the related simulation results. The benefits of applied charge recycling mechanisms

can be clearly seen by the output curve of Q . During each precharge phase the output nodes Q and \bar{Q} are not charged to V_{dd} but to significantly lower voltage of approximately $660mV$. This voltage is created after the equalizing effect of electron charges is balanced out between both outputs. Without any negative affection of the targeted voltage values during the evaluation phase (full swing range from $0V$ to V_{dd}), charge recycling leads to power savings of $\approx 34\%$, which is an estimable number. As a consequence, this power saving mechanism was integrated into the CR-DFF.

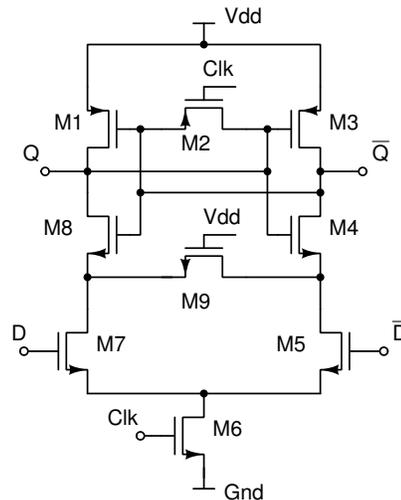


Fig. 16. CR Inverter

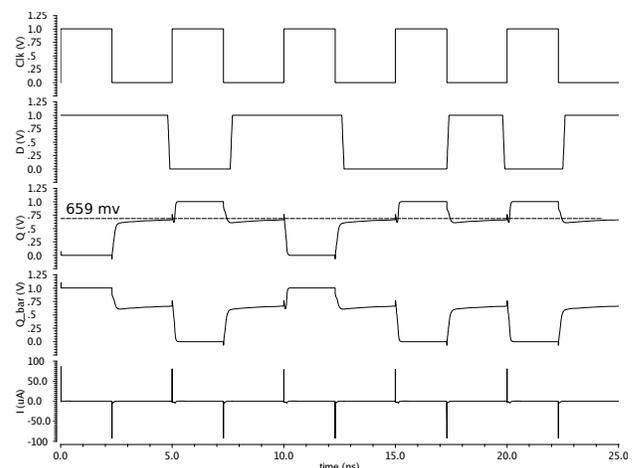


Fig. 17. Simulation results CR Inverter

The CR D-FF senses the inputs D & \bar{D} at the positive edge of Clk and stores these data independently from any changes at the input nodes of this circuit. Due to all implemented circuit improvements, an average static leakage current of $173nA$ is achieved, which is sufficiently low to be accepted. During the negative edge of Clk , the CR D-FF turns into the precharge phase, where all internal and external nodes are charged. The characteristic curves in Figure 18 show one beneficial features of the CR D-FF over the other discussed designs. This can be seen in both output curves of Q & \bar{Q} .

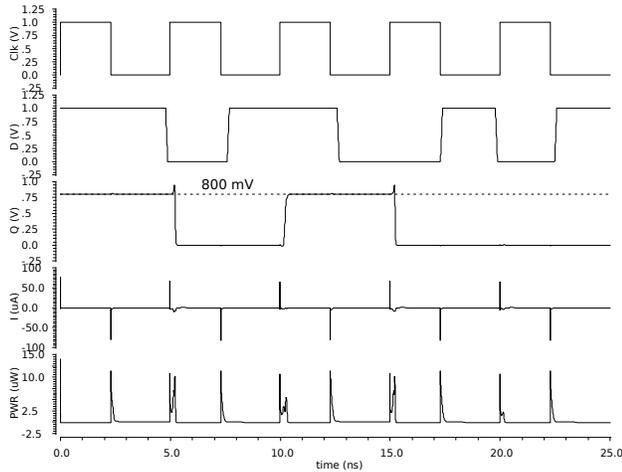


Fig. 18. Simulation Results CR D-FF

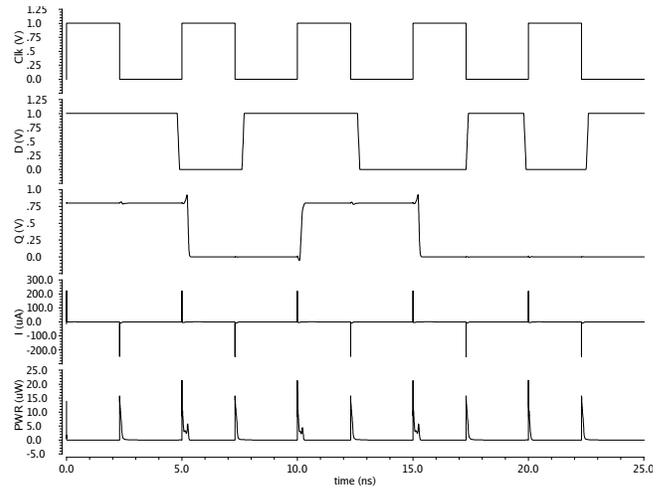


Fig. 20. Simulation Results SABL D-FF

The same simulation was applied with an implemented SABL D-FF, which is shown in Figure 19. The respective simulation results are also shown in Figure 20. Measuring the average power dissipation led to a result of $442.7nW$. A maximum power dissipation of $21.49uW$ and a minimum power dissipation of $22.73fW$ highlights the competitive results, which could be even better, especially when discussing about the average result. The maximum Δ in consumed power during a switching event is $\approx 26.17\%$, which is the second best result when compared to the selected designs.

results of the CR D-FF, shown in Table III.

TABLE III. SIMULATION RESULTS (PWR)

D-FF Type	Average PWR nW	Max. PWR uW	Min. PWR fW
Reference	1186	233.3	51.47
SETD	374.1	32.01	22.39
Low-power	275.7	73.89	19.92
PPI	435.4	110.5	172.3
CR	303.5	13.84	27.59

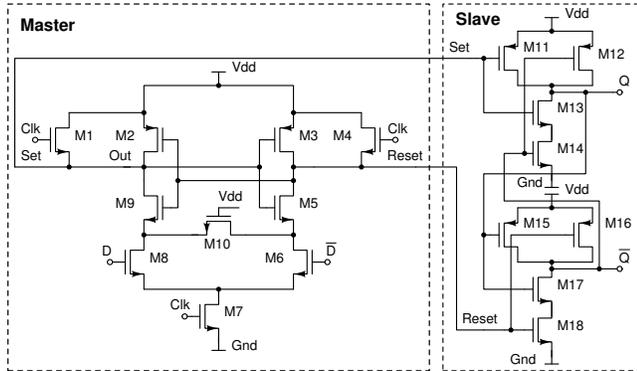


Fig. 19. SABL D-FF

Since this design features charge recycling, the output nodes and all internal nodes are precharged to $V_{dd} - V_{th}$ only, which is beneficial for the energy balance of this circuit. The reason for this is that precharge is finished by achieving an output voltage, which is one threshold voltage below V_{dd} . Thus, the less energy from the power supply is required for precharging the CR D-FF, the more suitable circuitry for low-power applications will be. Based on the reduced voltage range at the outputs of the master latch, it is possible to decrease permanently the supply voltage $V_{dd Slave}$. Hence, we choose a supply voltage of $800mV$ for the conventional slave circuit, which supports further power dissipation reduction. For a better comparison, we enhance Table I with relevant simulation

The results in Table III show that the introduced CR D-FF outperforms most of the previously analyzed designs in terms of average power consumption. It achieves the second-best performance for average power consumption ($319.7nW$) and the best result for maximum power dissipation ($13.84uW$). The minimum power consumption of $27.59fW$ can be neglected, since the influence of these contributions is not significant for the overall performance of all discussed designs. Even though the conventional low-power flip-flop achieves a slightly lower average power consumption than the CR D-FF, the peak power dissipation is approximately quintuple higher and it offers no resistance features against DPA. Figure 21 shows a comparison of the average power consumption. Additionally, the differences in average power consumption are highlighted in Figure 22.

It can be clearly seen in Table III that the CR D-FF provides the most constant power consumption among all considered designs, therefore also providing the best opportunities to be chosen in security-sensitive applications. The smaller the differences in energy consumption between each data transition are, the more difficult a differential power analysis will be, which is always the starting point for a side-channel attack. Hence, the introduced CR D-FF provides both, remarkable low-power characteristics for mobile, embedded circuitry, which comes along with a necessity for robustness against intended attacks. However, benefits in superior energy efficiency and noticeable robustness against differential power analysis come at the cost of a higher number of transistors, shown in Table IV and in Figure 23.

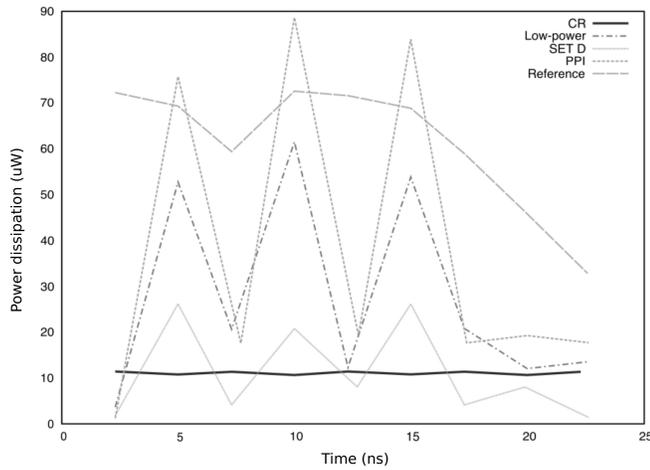


Fig. 21. Comparison of Average Power Dissipation

TABLE IV. TRANSISTOR COUNT AND POWER VARIATION

D-FF Type	Reference	SETD	LP	PPI	CR
No. of transistors	16	10	16	18	21
Max. PWR Δ (%)	18.78	94.7	94.03	98.62	6.8

This fact usually leads to a penalty in required area for manufacturing, which is certainly an aspect to be considered. A CR D-FF consists of 21 transistors and requires preparation of complementary input signals, which depend on additional wiring and therefore lead to extra area on the chip. On the other hand, this implementation provides also 2 complementary outputs with no delay between both signals and no necessity of additional circuitry for generation. Table IV also emphasizes the differences between the analyzed cells in switching behavior. Whilst the Δ of dissipated power of the CR D-FF never exceeds variations of 6.8% in maximum, the results of the alternative designs show much higher noticeable differences. Despite the fact that all designs have been analyzed without putting a stronger focus on speed and timing aspects, further measurements on the maximum operating frequency have been done. For this purpose, the

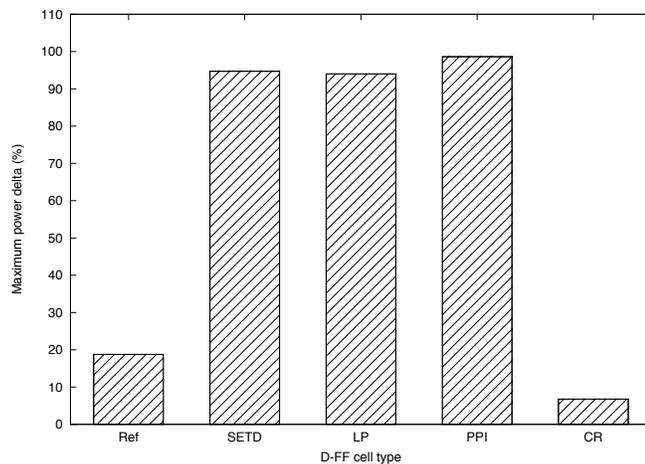
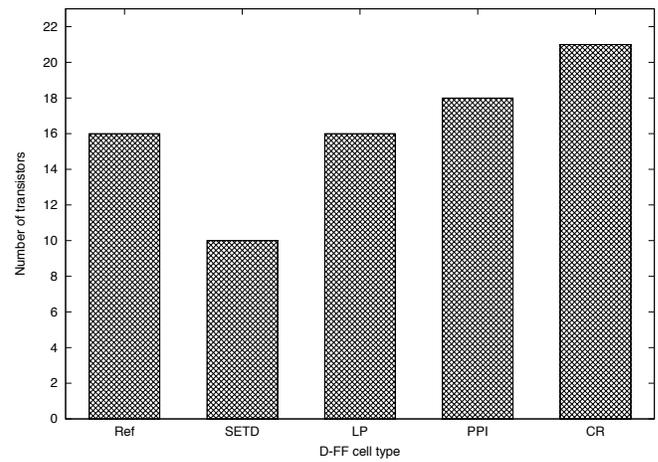

 Fig. 22. Measured PWR Δ


Fig. 23. Differences in Transistor Count

elapsed time for each switching transition was measured and compared against each other. Figure 24 illustrates a direct comparison of the output Q of all considered circuits after being stimulated with an input signal D . Depending on the switching transition and the characteristics of the flip-flops, expected differences on the edge steepness can be identified.

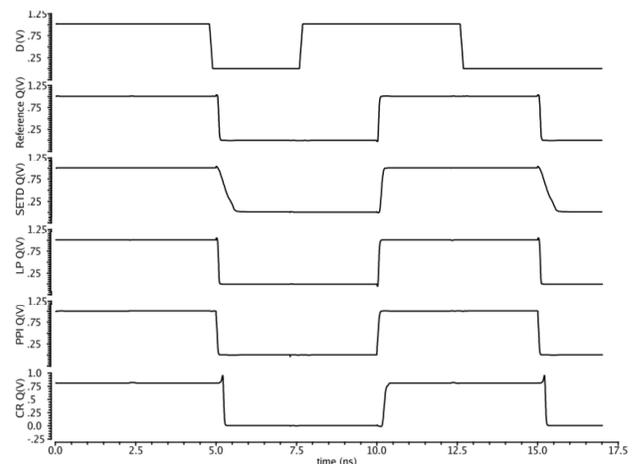


Fig. 24. Comparison of Switching Transitions Of All Designs

Based on these simulation results, the consumed time for a $HIGH \rightarrow LOW$ and a $LOW \rightarrow HIGH$ transition has been measured and summarized in Table V. For the sake of a better overview, these results are additionally illustrated in Figure 25 and Figure 26. The maximum achievable switching frequency f_{max} , which is illustrated in Figure 27 as a comparative overview, reveals the penalty in operating speed of the CR D-FF, due to the increased number of transistors. However, a maximum switching frequency of $\approx 6.4GHz$ is still a notable result. It shall be mentioned that even better results in terms of speed could be achieved by a further fine tuning of the transistor parameters. Especially p-channel transistors may be a bottleneck when it comes to circuit's speed optimization.

TABLE V. TIMING COMPARISON

D-FF Type	T High-Low ps	T Low-High ps	Max. freq. GHz
Reference	42.5	58.3	9.9
SETD	422	101	1.9
Low-power	43.63	51.58	10
PPI	60.48	79.16	7.1
CR	41	114	6.4

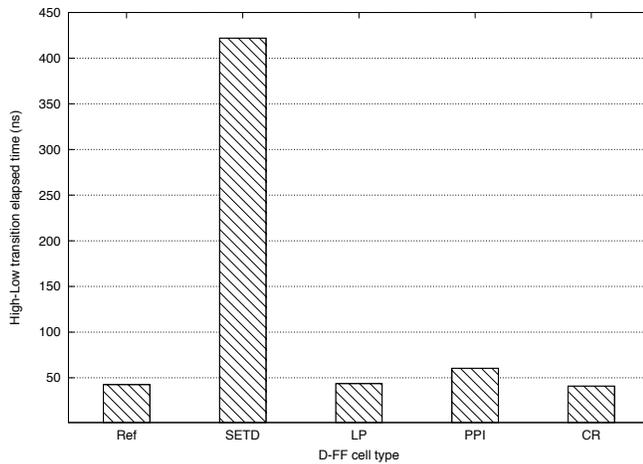


Fig. 25. Transition Time For HIGH to LOW Switching Event

V. CONCLUSION

We analyzed a selected number of existing flip-flop designs upon their characteristics and suitability for usage in low-power applications. Beside that, we have investigated each design on its capabilities to be resistant against differential power analysis. Our goal was to design a D-FF, which provides both, a remarkable reduction of power consumption and robustness against side-channel attacks. An intermediate step towards our final circuit was the implementation of the SABL FF, which can be considered as a predecessor to our intended design. It is a dynamic and differential logic with two different, altering operational modes. Simulations of the SABL FF have proven that its capabilities in terms of DPA resistance are consider-

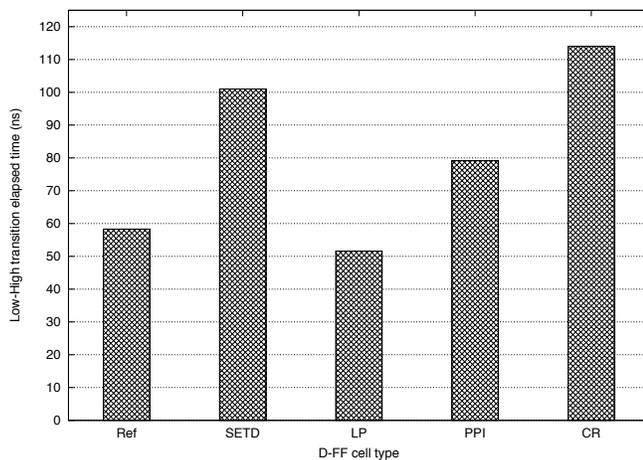


Fig. 26. Transition Time For LOW to HIGH Switching Event

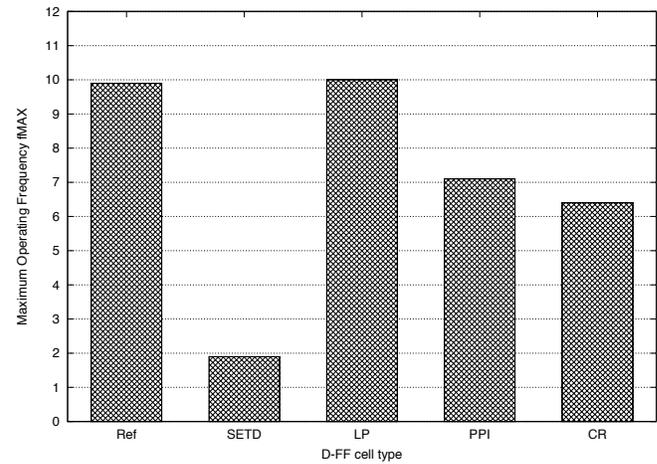


Fig. 27. Comparison Of Maximum Operating Frequencies

able, but its low power characteristics could be still improved by trying to re-use internal electric charges for support of the power supply during *precharge*. Hence, we designed a charge recycling D-FF, which uses not discharged electrons at one of the complementary output nodes to support the battery during the precharge phase. This benefit comes along with the fact that the outputs of the master latch are precharged to $V_{dd} - V_{th}$ only, providing the opportunity to power the slave latch with the same supply voltage ($\approx 800mV$). Furthermore, we applied additional power saving modifications and achieved substantial improvements of power reduction and standby leakage suppression. Simulation results have shown that the CR D-FF offers the best overall performance with an average power consumption, which reduced the dissipated power by about $\approx 75\%$. Complementary generation of output signals with no requirement for delay correction is a further advantage of this circuit when compared to other designs, which do not feature parallel, complementary creation of D & \bar{D} . The variations of the measured power consumption do not exceed differences of $\approx 7\%$ and remain constant independent from the switching event, which is sufficient to withstand differential power analysis and is not achieved by the alternative flip-flops. These benefits come at the cost of a higher number of required transistors and the layout after synthesis of a CR D-FF requires careful routing of all metal interconnections between these cells for keeping the parasitic capacitances as equal as possible. Another drawback could be the necessity for provision of complementary inputs, leading to additional inverters. All the results were achieved with out-of-the-box transistor parameters, since our intention was to investigate whether acceptable results could be achieved without modifications to width or even length of each transistor. Next steps can be carried out to further improve the circuit's attributes and one of them could be a detailed analysis of a transistor's fine tuning impact on the overall performance. Nevertheless, each enlargement of silicon area can result in asymmetric wiring and therefore in a penalty of DPA robustness.

ACKNOWLEDGMENT

The authors thank Pierre Mayr, from Ruhr University of Bochum, for his advice on verification strategies and procedures. We would like to give credit to Grant Martin, from Cadence Tensilica, for many interesting discussion about embedded devices and low-power technologies. We are grateful to Andreas Ullrich, from University of Wuppertal, for immediate PDK / tool support.

REFERENCES

- [1] K. Niewiadomski and D. Tutsch, "Low power charge recycling D-FF," in *Proceedings of the 10th International Conference on Advances in Circuits, Electronics and Micro-electronics (CENICS 2017)*, September 2017, pp. 21–27.
- [2] S. Fürst, "Challenges in the design of automotive software," in *Proceedings of the Conference on Design, Automation and Test in Europe*, ser. DATE '10. 3001 Leuven, Belgium, Belgium: European Design and Automation Association, 2010, pp. 256–258, last accessed on 2018-05-29. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1870926.1870987>
- [3] M. Ullmann, M. Hübner, B. Grimm, and J. Becker, "An FPGA run-time system for dynamical on-demand reconfiguration," in *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International. IEEE*, 2004, p. 135.
- [4] R. A. et al., "Towards a dynamically reconfigurable automotive control system architecture," in *Embedded System Design: Topics, Techniques and Trends*. Springer, 20017, pp. 71–84.
- [5] K. Niewiadomski, C. Gremzow, and D. Tutsch, "4T loadless SRAMs for low power FPGA LUT optimization," in *Proceedings of the 9th International Conference on Adaptive and Self-Adaptive Systems and Applications (ADAPTIVE 2017)*, February 2017, pp. 1–7.
- [6] K. Niewiadomski and D. Tutsch, "Enhanced 4T loadless SRAM comparison with selected volatile memory cells," in *International Journal on Advances in Systems and Measurements (IARIA)*, December 2017, pp. 139–149.
- [7] V. Stojanovic and V. G. Oklobdzija, "Comparative analysis of master-slave latches and flip-flops for high-performance and low-power systems," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 4, pp. 536–548, Apr 1999.
- [8] K. Tiri and I. Verbauwhede, "A VLSI design flow for secure side-channel attack resistant ICs," in *Proceedings of the Conference on Design, Automation and Test in Europe - Volume 3*, ser. DATE '05. Washington, DC, USA: IEEE Computer Society, 2005, pp. 58–63, last accessed on 2018-04-16. [Online]. Available: <http://dx.doi.org/10.1109/DATE.2005.44>
- [9] *XA Spartan-3A Automotive FPGA Family Data Sheet*, Xilinx, 04 2011, rev. 2.0.
- [10] R. E. Aly, M. I. Faisal, and M. A. Bayoumi, "Novel 7T SRAM cell for low power cache design," in *Proceedings 2005 IEEE International SOC Conference*, Sept 2005, pp. 171–174.
- [11] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital integrated circuits- A design perspective*, 2nd ed. Prentice Hall, 2004.
- [12] C. Maxfield, *The Design Warrior's Guide to FPGAs: Devices, Tools and Flows*, 1st ed. Newton, MA, USA: Newnes, 2004.
- [13] M. Sharma, A. Noor, S. C. Tiwari, and K. Singh, "An area and power efficient design of single edge triggered D-Flip Flop," in *2009 International Conference on Advances in Recent Technologies in Communication and Computing*, Oct 2009, pp. 478–481.
- [14] U. Ko and P. T. Balsara, "High-performance energy-efficient D-flip-flop circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 1, pp. 94–98, Feb 2000.

Capillary Sensors with UV-Forced Degradation and Fluorescence Reading of Chemical Stability and Polycyclic Aromatic Hydrocarbons Presence in Diesel Fuels

Michał Borecki

Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Warsaw, Poland
e-mail: borecki@imio.pw.edu.pl

Michael L Korwin -Pawlowski

Département d'informatique et d'ingénierie
Université du Québec en Outaouais
Gatineau, Québec, Canada
email: michael.korwin-pawlowski@uqo.ca

Mateusz Gęca

Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Warsaw, Poland
e-mail: mati.geca@gmail.com

Przemysław Prus

Blue Oak Inventions
Wrocław, Poland
email: pprus@boinv.com

Jan Szmidt

Institute of Microelectronics and Optoelectronics
Warsaw University of Technology
Warsaw, Poland
e-mail: j.szmidt@imio.pw.edu.pl

Abstract - There are many standards and types of laboratory equipment for examination of specific properties of diesel fuels. The basic standards of diesel fuel stability require examination taking a relatively long time, counted in days. The development of new methods of diesel fuel stability testing has as its aim accelerated ageing of the examined samples. The most popular of accelerated ageing factors are the increase of temperature and oxidation. In this paper a new method with diesel fuel sample positioned in a capillary and ultraviolet radiation used as degradation factor is proposed. Two possible optical sensor configurations are described as well as the data analysis method for classification of premium and standard commercial diesel samples. The comparison of two sensor configurations was made with the same fuel samples which included winter and summer premium diesel fuel as well as eco winter diesel fuel, unmodified and modified with sludge protection additive. The results of sensor analysis during fuel examination prove that 40 minutes of UV degradation and sequential fluorescence reading at 10 selected moments of time coupled with data processing is enough to evaluate diesel fuel chemical stability and quality. In the experiments light of 265nm and 365nm wavelengths was used correspondingly for degradation and fluorescence reading. We found that chemical stability of fuels was related to the amplitude variations of characteristic emitted fluorescence signals. The concentration of polycyclic aromatic hydrocarbons in fuels was related to the amplitude of signals emitted from excited samples. The UV examination indicated that fuel's chemical stability was better observable with forced degradation and excitation at 265nm, while fuel's polycyclic aromatic hydrocarbons presence was better observable with excitation at 365nm.

Keywords-diesel fuel; fuel quality; fuel stability; PAHs detection, capillary sensor; fluorescence sensor; UV fuel degradation.

I. INTRODUCTION

This paper's focus is on selected aspects of a sensor using a capillary optrode and UV radiation as the fluorescence excitation factor and fuel degradation initiator that enables rapid testing of chemical diesel and biodiesel fuels stability. The principle and preliminary results of this work were presented in [1].

A. Diesel and biodiesel fuel –as seen from the user's and producer's sides

On the diesel fuel user's side low cost and high quality of the fuel are important. Diesel fuel quality is analyzed with reference to diesel engine operating characteristics, such as starting ease, low noise, low wear, long life, sufficient power, low temperature operability, and low emissions [2].

On the fuel producer's side, the quality of diesel fuel is determined by quality standards. These quality standards differ for diesel and biodiesel fuels. The most popular standards of diesel fuels are ASTM D 975 introduced by the American Society for Testing and Materials and EN 590 introduced by the European Committee for Standardization. Respectively, the standards for biodiesel fuels are the ASTM D 6751 in the USA and the EN 14214 in Europe. In this approach, the quality of fuel can be described as a set of laboratory measured fuel parameters, like for example: the

cetane number, pour point, viscosity, density, acid number, the result of accelerated oxidation stability test, and result of test for gum content in fuels by jet evaporation, as well as sulfur, nitrogen and oxygen contents of fuel composition.

The common feature of fuel user's and producer's sets of fuel quality parameters is that they depends on diesel fuel composition, which changes in storage due to presence of some chemical active fuel components [3].

B. Type and composition of diesel and biodiesel fuels

The diesel and biodiesel fuel consists of the fuel base, the fuel improvers and impurities. The type of diesel fuel may be described by the fuel base origin or the technological process.

Petroleum diesel (petrodiesel) is the fuel produced in a refinery as a blend of straight-run product, fluid catalytic cracking light cycle oil, and hydrocracked oil. The straight-run product may be acceptable as diesel fuel. Synthetic diesel (syndiesel) can be produced from any carbonaceous material. The raw material is gasified and after purification is converted in the Fischer-Tropsch process. The synthetic diesel generally is the first class fuel base. Fatty-acid methyl ester (FAME), obtained from vegetable oil or animal fats which have been transesterified with methanol is a common biodiesel fuel base composition. FAME has lower oxidation stability than petrodiesel, and it offers favorable conditions for bacterial growth. Hydrogenated oils (HVO) may be produced from the triglycerides from vegetable oil and animal fats. The HVO is a composition of alkanes obtained in the refining and hydrogenation process. Therefore, HVO composition is similar to synthetic diesel like fuel.

Modern petroleum-derived diesel fuel base is composed of about 74% saturated hydrocarbons - primarily alkanes, 25% aromatic and acyclic unsaturated hydrocarbons. The typical chemical formula for diesel fuel base molecule composition is $C_{12}H_{23}$, but the number of carbons in the molecule ranges approximately from C10 to C22. Now, it is worth noting here, that at 25°C alkanes ranging from C5 to C17 are liquids, while alkanes ranging above C17 are in solid state. Therefore, it is expected that in the diesel fuel some solid state hydrocarbons particles are present and form some kind of colloid.

The most common saturated hydrocarbon of diesel fuel is dodecane described with the formula $H_3C(CH_2)_{10}CH_3$. Alkanes have the general chemical formula C_nH_{2n+2} . Alkanes are only weakly reactive with ionic substances, but all alkanes react with oxygen in a combustion reaction [4].

Saturated cyclic liquid hydrocarbons are not alkanes, but they are often called cycloalkanes as are characterized with similar chemical properties as alkanes.

Unsaturated acyclic hydrocarbons fuel content consist primary of alkenes, but alkynes, alkadienes and alkatrienes may be found in diesel fuel.

Acyclic alkenes, with only one double bond and no other functional groups are described by the general formula C_nH_{2n} . It can be seen that for the same number of carbon atoms acyclic alkenes have two hydrogen atoms less than the corresponding alkane. Acyclic alkyne is an unsaturated hydrocarbon containing at least one carbon-carbon triple

bond. The general chemical formula of basic alkyne is C_nH_{2n-2} .

Alkadienes are acyclic hydrocarbons having two carbon-carbon double bonds, while alkatrienes are hydrocarbons having three carbon-carbon double bonds. A simple alkadiene is butadiene, which is a gas with the structure $H_2C=CH-CH=CH_2$. This molecule is important in the chemical industry as is used in production of synthetic rubber. A popular acyclic hydrocarbon containing three carbon-carbon double bonds; is 2,4,6-octatriene, which has the structure $H_3C-CH=CH-CH=CH-CH=CH-CH_3$ and a standard name isooctane. It is a liquid and an important component of gasoline used in relatively large proportions to increase the knock resistance of the fuel.

Unsaturated aromatic (cyclic) hydrocarbons are represented by monocyclic benzene (MAH - monocyclic aromatic hydrocarbon) and by polycyclic aromatic hydrocarbons (PAHs). Benzene is a natural liquid constituent of crude oil and is one of the elementary petrochemicals. The simplest PAH having two aromatic rings is naphthalene (white crystalline solid). The three-ring PAHs compounds are anthracene (colorless solid) and phenanthrene (white powder).

PAHs molecules can be found in coal and in tar as well as in crude oil. Therefore, almost every petroleum product contains various PAHs particles [5]. Diesel fuel is expected to contain PAHs particles ranging from three to five ring systems.

Fuels fabricated for diesel engine may contain bio-component that may reach a few percent of fuel volume. The bio-components' presence is significant from fuel technology, affiliation (bio-diesel) and fuel parameters points of view [6]. Important is that bio-components may be of first or second generation [7]. The first generation of bio-diesel components includes fatty acids methyl esters (FAME) that are characterized by the presence of double bonds, while the second generation components base on hydro-treated vegetable oils (HVO) that seem to be mainly carbohydrates with saturated carbon atoms bonding [8].

The modern diesel fuels include also a range of additives, as for example cetane improvers, antioxidants, detergents, corrosion inhibitors, deposit modifiers, lubricity agents, and biocides [9].

The most popular cetane improver is 2-ethyl hexyl nitrate (2-EHN); the minor significant improver is di-tertiary-butyl peroxide (DTBP). DTBP is a peroxide and its chemical formula is $(CH_3)_3COOC(CH_3)_3$. 2-EHN is nitroalkane, which linear formula is $CH_3(CH_2)_3CH(C_2H_5)CH_2ONO_2$. Both substances include oxygen and are reactive.

It was found that hydrocarbons mixture with C10-C18 particle is the optimum range for microbial metabolism [10]. Therefore, in stored diesel fuel some microorganisms are presented. The microorganisms adhere to the tank walls or settle in the fuel-water interface at the bottom of the tank. Possible routes of entry of microorganisms into the fuel can be from the ground by cracks of buried tanks and from the air, during loading operations. Furthermore, organic biocides with different functional active groups and hydrogen peroxide are in use as additives to diesel fuel [11].

C. Diesel fuel degradation during storage process

Hydrocarbons of diesel fuel may degrade as a result of oxygen presence and reaction initiator. Known reaction initiators are light, temperature and catalyst.

The photochemical reactions caused by light are mainly able to affect the physical properties of some of the oil fractions. They are able to alter the emulsion formation and the solubility of the fuel fractions. This is done by inducing reaction between oil components and other molecules, which makes the molecules more polar and water soluble, so creating new compounds with other physical properties [12]. It seems that there major mechanism involved in photo degradation of oil can be classified as direct or indirect photolysis [13]. Direct photolysis takes place when the molecule of interest absorbs energy from light and further degrades. Indirect photolysis occurs when another molecule absorbs the energy from light and reacts to degrade the molecule of interest. The UV radiation also can suppress a number of microorganisms in diesel fuel [14].

The thermal stability of diesel fuels was examined at different temperatures: higher than 150°C and below 150°C [15].

Saturated hydrocarbons common degradation is a result of oxidation. The products of oxidation include aldehydes, ketones, acids and alcohols which can further transform into multi-particle substances that form resins and sediments [16].

The unsaturated hydrocarbons degradation in standard vessel condition may occur as the oxidation and polymerization process. Certain alkenes undergo self-addition reactions in the presence of specific catalysts or energy supply to produce molecules called polymers. The reaction involves double bonds being converted to single bonds as hundreds of molecules bond and form long chains. It is not possible to give an exact formula for a polymer produced by a polymerization reaction because the individual polymer molecules vary in size [17]. Polymers made from alkenes result in a very long-chain alkane. As a result, it has the chemical inertness of alkanes and form solid sediment in fuel.

Chemical oxidation of PAHs occurs readily, for example, anthracene (C₁₄H₁₀) oxidation gives anthraquinone, (C₁₄H₈O₂). PAHs react also with singlet oxygen - a high-energy form of oxygen. PAHs react with acetone in photo degradation. For example, naphthalene with 1% acetone mixture during UV irradiation results in 1,2-Benzenedicarboxaldehyde generation. The examination at 254, 310 and 365 nm, wavelength shows that the influence of wavelengths on photodegradation rates of PAHs was significant [18]. The irradiation at 365nm found to be effective in oxidation of PAHs. For example, naphthalene C₁₀H₈ intermediates produced during UV irradiation in oxygen presence are:

- C₈H₆O₂ - 1,2-Benzenedicarboxaldehyde,
- C₉H₆O₂ - 2H-1-benzopyran-2-one,
- C₁₀H₆O₂ - 1,4-Naphthalenedione,
- C₁₀H₈O - 1-Naphthalenol,
- C₁₀H₆O₃ - 2-Hydroxy-1,4-naphthalenedione, [19].

Cetane improvers decompose rapidly and form free radicals when exposed to temperatures above 100°C. These radicals increase the rate of main fuel components decomposition; therefore, the ignition delay is negatively affected [20].

DTBP molecule during decomposition supplies both the oxidizer and the fuel particles. Thus, it can be used as a fuel in engines where oxygen is limited [21].

2-EHN in liquid phase can decompose in the absence of air as, for example, when pressure and temperature in storage tank increase above 75°C [22]. But, in another report [23], authors suggest that at temperatures below 100°C the decomposition of 2-EHN in diesel fuel is extremely slow, as at 120°C the half-life is 522.7 h. Moreover, in the mentioned paper, the 2-EHN at the 146.3°C and 2400 bar half-life is 88.5 h, while half-life at 1 bar is 16 h. The initial step of 2-EHN particle decomposition is the separation of NO₂ and the rather unstable 2-ethylhexyloxy radical (2-EHO). Thus, we conclude that 2-EHN behavior in diesel fuel environment may depend on fuel composition. Moreover, such conclusion is supported by proposed method of 2-EHN concentration examination in diesel fuel with UV fluorescence examination reported in [24].

The antioxidant additive of diesel fuel principal function is to inhibit the formation of peroxides. Therefore, as can be seen, antioxidants should react with some cetane improvers as for example DTBP.

D. Sensing methods for diesel fuel quality

There are many methods to measure diesel fuel quality. The standardization process of measurement of diesel fuel parameters is still in progress [25]. Nowadays, the basic idea of fuel characterization is a selective and specific measurement of sequent parameters.

Petro-chemic industry representative recognized a series of fuel parameters measured in laboratory conditions, as for example a cetane number (CN), acid-number, flashpoint, lubricity and oxidation stability. Measurements of the mostly recognized fuel quality parameter called ignition quality of fuel (CN) have to be carried out in the Cooperative Fuel Research Engine (CFR-5) or the Ignition Quality Tester (IQTTM). On the other hand, many fuel parameters depend on fuel composition described by distillation results. The distillation parameters may to be measured using the classic distillation unit, or may be measured with automated distillation process analyzer [26]. Oxidation stability of fuel standard measurement with rancimat method refers to oils chemical reaction with gas environment [27]. It should be noted that chemical stability of fuel means thermodynamic stability of a chemical system and refers to chemical internal reaction between fuel components, [28] [29]. Therefore, chemical stability of fuel is not the same as oxidation stability. But, some definitions of the chemical stability refers to the tendency of a material to resist change or decomposition in its natural environment when exposed to air, heat, light, pressure, or other natural conditions.

Contrariwise, fuel users expect parameters set that describe diesel fuel operability and that can be measured in situ. Unfortunately, such set of parameters is not full

standardized, but include engine start characteristic, time of fuel life (stability), fuel transfer dynamics from tank to injection unit and fuel contamination.

A set of method for fuel contamination with different chemicals as for example non-transferred vegetable oils, kerosene, PAHs are under development. The examination of bio-components concentration can be made with the use of excitation-emission matrix method for wavelength range of 250-500nm with spectral subtraction fluorescence for range of 300-600nm. This method is precise, but requires expensive equipment [30]. Therefore, mentioned method has been adopted to use with the one light emitted diode as excitation source [31]. But, an investigation result refers only to samples of diesel fuel intentionally contaminated with known concentration of bio-component or kerosene. It is worth noted that time-resolved fluorescence method can be used for classification of the unknown fuels groups as gasoline, diesel, bio-ester and vegetable oil [32]. The fluorescence time delay method and a high pass method are used for the detection of PAHs. In the high pass method, the high pass optical filter was used to remove the scattering of the exciting pulse, as the exciting energy may be too high and destroy detector. For example, naphthalene detection was performed with excitation at wavelength of 266 nm, with 6 ns pulse of 5 mJ, which was repeated at 10 Hz rate [33].

A set of method oriented to examination of oxidation stability of diesel fuels are also under investigations. For example, the method of antioxidant presence in fuel samples with absorption measurement performed at 256 nm has been proposed [34].

Synchronous fluorescence spectroscopy was suggested as a tool to evaluate the photo and thermal degradation results of crude oils [35]. The degradation was performed with sunlight of 350 W/m² that lasts for 100 hours. Oils irradiated showed decreased fluorescence intensity at all wavelengths from 350 to 700nm, when examined in intervals of at least 2 hours.

New methods of diesel fuel fitness-for-use determination that are addressed to implementation in sensors are under development. These methods include measurement of characteristic points of local sample heating of fuel positioned in a capillary [36], measurement of the dynamical rise of fuel in an inclined capillary [37], reading of fluorescence effects in the range of UV-VIS and reading of UV-forced degradation results of closed sample of diesel fuel [38]. Dedicated components of mentioned sensors, as capillary optrodes [39], micro-heaters [40], optoelectronic photodetectors [41] and photodetectors pre-amplifiers [42], have been developed. Also, some aspects of automatization of sensor with capillary set-ups as data filtering and processing [43] as well as influence of man operator are under investigation [44]. Moreover, progress in explored methods of diesel fuel examination is coupled with development of new photonics elements as fluorescent core micro-capillaries [45], improvements in construction of electronic components as electronic modules [46] [47] and improvement in microelectronic manufacturing technology [48].

E. Photo degradation and photo detection possibilities of potentially unstable diesel fuel components

The photo degradation and photo detection of potentially unstable fuels components is possible with radiation that wavelength range is absorbed by fuel in particular by unstable particles.

The maximum absorption of alkenes shifts from deep UV toward longer wavelengths with the increase of the number of carbon-carbon double bonds, for example, H₂C=CH₂ the maximum absorption is at 180nm, for H₂C=CH-CH=CH₂ the maximum absorption is at 217nm, while B-carotene with 11 double carbon-carbon bonds is characterized by the maximum absorption at 455nm [49]. Absorption of UV in oil samples at 230–235 nm allows detection of alkadienes, while absorption at 260, 270 and 280 nm allows detection of alkatrienes [50]. For example, detection of isooctane may be performed as absorbance measuring at 232 nm and 264-272 nm.

Similar dependency of maximum absorption shift can be observed in aromatic unsaturated hydrocarbons. For example, benzene maximum absorption is at 255nm, naphthalene is at 286nm while anthracene is at 363nm. Absorption peaks are also influenced by functional groups while benzene maximum absorption is at 255nm, phenol is at 270nm, and nitrophenol is at 320nm.

PAHs generally absorb light in the 200-400 nm range and strongly fluoresce in 360-430nm range. For example, peak wavelength of excited fluorescence for benzo[a] anthracene is 387nm, for benzo[a]pyrene is 405nm, for benzo[k]fluoranthene is 403nm, for chrysene is 382nm and for phenanthrene is 366nm [51]. It is interesting that the emission spectra of benzo[a]anthracene is characteristic and can be described as hills with two similar peaks with one a little dominant height at 387nm and 410nm and with wavelength symmetric saddle. The naphthalene also has a similar shape with hills peak at 325nm and 335nm. The photo degradation of naphthalene with the addition of nitrogenous air pollutants NO₂ was described [52]. The results indicate that naphthalene degradation occurs when irradiated with UV light > 290 nm. Reaction products were 2-formyl-cinnamaldehyde, 1,4-naphthoquinone, nitronaphthol, o-phthaldialdehyde, phthalide and nitronaphthalene. Nitronaphthalene is characterized by unique solvent-dependent fluorescence [53].

Up today, PAHs photo degradation as pyrene and benzo[a]pyrene using ultraviolet sources working at wavelengths of 254, 310 and 365 nm was investigated. The results of UV irradiation intensity on the degradation performance of PAHs show that with the intensity of UV irradiation being increased, the degradation rates of PAHs were higher. Additionally, degradation rates of pyrene were different at different UV irradiation wavelengths [54]. Also, PAHs degradation in real outdoor condition showed a good linear correlation with mean solar radiation intensity [55]. But, it was indicated that photo degradation rates of anthracene, phenanthrene and naphthalene in water surfaces are different. Anthracene half-life was 1 hour, phenanthrene

was 20.4 hours while half-life of naphthalene exceeded 100 hours [56].

F. Intermediate conclusions – selection of UV sources to degradation of unstable fuel components

The diesel fuel is a multicomponent mixture that may be degraded with oxygen, temperature or light irradiation. Classical methods of fuel analysis of oxidation stability require trained chemist, costly equipment and prolonged time of examination or are not precise. On the other hand, fluorescence is a sensitive indicator of photochemical transformations of fuel components [57]. Therefore, examination of chemical stability of diesel fuels may be performed with forced UV degradation and fluorescence reading of signal emitted from unsaturated hydrocarbons.

It should be noted that light absorption in popular PAHs as naphthalene {288 nm}, phenanthrene {252 nm; 275 nm; 294 nm} and anthracene {252 nm; 356 nm} can be addressed with 260 UVTOP LED that emits radiation from 255 nm to 270 nm when calculated for 20% of peak intensity located at 265 nm. This is due fact that aromatic components light absorption characteristics are not narrow peaks of mentioned wavelengths but extends into quite wide bands. For example, naphthalene absorption calculated at half intensity extends from 250nm to 289nm [58].

On the other, hand solar radiation falling to the ground does not include UVC range, but include UVA. The example of intensity of spectral distribution of sun irradiation for sunny day of 09 of May 2018 in Warsaw is presented in Figure 1.

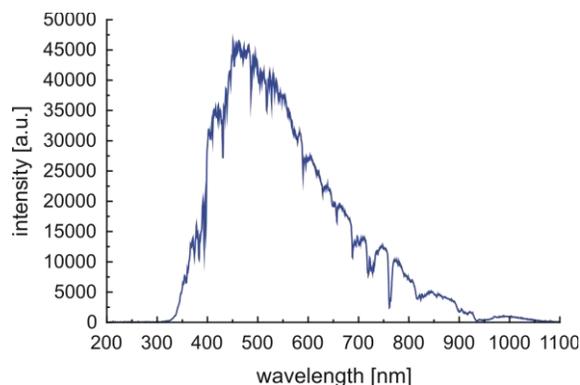


Figure 1. Distribution of intensity of sun irradiation in sunny day of 09 of may 2018 in Warsaw registered at 4pm.

This data may lead to conclusion that radiation of 365nm wavelength is well represented in natural environment. Besides, for such radiation high powers LEDs are available, as for example M365F1.

Therefore, the rest of this paper is organized as follows. Section II describes the sensor construction development including sensor head with optrode as well as optoelectronics system set-up. Section III addresses the results of the analysis fluorescent signal that is recorded during photo degradation at selected wavelength. Section IV goes into conclusions.

II. SENSOR CONFIGURATION

Proposed capillary sensor with UV-forced degradation and fluorescence reading of chemical stability of diesel fuels can be configured with one, two or three UV sources. The UV forced degradation and fluorescence excitation in sensor with one UV source is performed at the same wavelength. The degradation and fluorescence excitation wavelengths can be different in sensor with two or three UV sources. The sensor configuration with a couple UV sources gives possibility to spread wavelengths for degradation and for fluorescence excitation, but the number of combination of measurement configuration increase rapidly with sources number. Therefore, in this paper the configurations of sensors with forced degradation and fluorescence excitation at the same wavelength are examined. The selected wavelengths for sensor examination are 265nm and 365nm. As a UV sources the light emitting diodes are in use correspondingly UVTOP265 LED and M365F1 LED. The examined UV degradation is realized by UV irradiation of sample fuel positioned in capillary for 40 minutes. During this degradation, fluorescence signals are read at 0.5, 1, 5, 10, 15, 20, 25, 30, 35 and 40 minutes.

A. Sensor set-up with one UV source working at 265 nm

The sensor head of the set-up with one UV source working at 265 nm is presented in Figure 2.

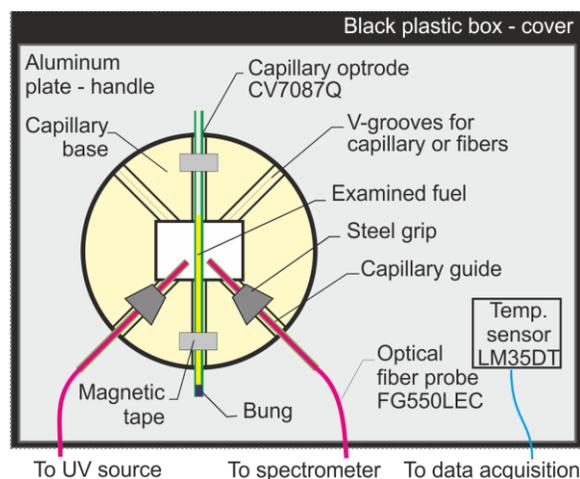


Figure 2. Scheme of the capillary sensor head with one UV source.

The head is coupled to the optoelectronic sensor set-up for UVTOP265 LED according to the scheme presented in Figure 3. The set-up is optically powered by a LED controlled with the hardware D2100 driver that is triggered from a PC with the use of software. The fiber optic divider 1:7 is used to produce a reference signal and to monitor light source parameters with PM100D optical power meter equipped with S120VC photodiode. The head output is connected to Maya 2000pro spectrometer, which is connected to PC and controlled by Ocean Optics software that enables sequential writing of data recorded.

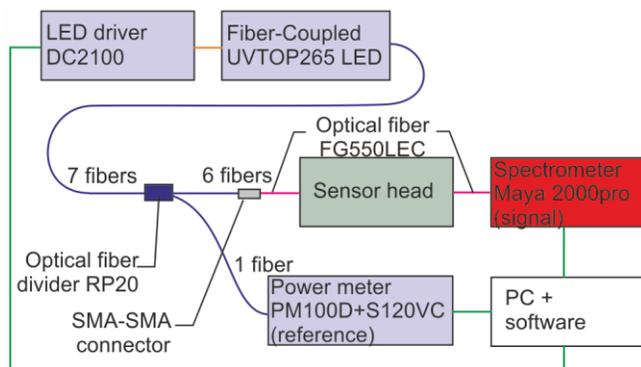


Figure 3. Scheme of the capillary sensor set-up with one UVTOP265 LED.

Data recorded from spectrometer and power meter are processed in each examination for elimination of spectrometer dark signal, light source emitted power drift reduction and for de-noising of signal.

B. Sensor set-up with two UV sources working at 365nm

The sensor set up for head working at 365 nm requires relatively narrow band for fluorescence excitation source as the wavelength of excitation signal is close to emitted band of some PAHs. For example, M365F1 LED bandwidth (FWHM) is 7.5nm and the residual long wavelength radiation may interfere with excited fluorescence signals. Therefore, M365F1 LED using as fluorescence excitation requires optical filtering, which may reduce optical power required for degradation. Thus, dividing optical channel to excitation and degradation leads to the sensor configuration proposed in Figure 4. The head is coupled to the optoelectronic sensor set-up with M365F1 LEDs according to the scheme presented in Figure 5. The sensor set-up for fluorescence excitation and reading with M365F1 LED, contrary to UVTOP265 LED, includes a monochromator adopted as precise optical filter.

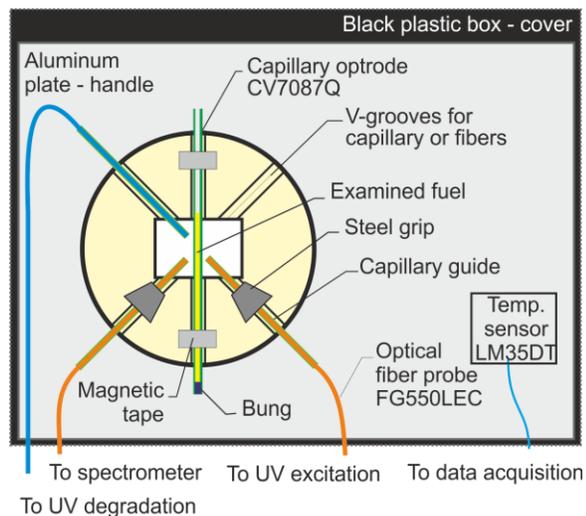


Figure 4. Scheme of the capillary sensor head with two UV sources.

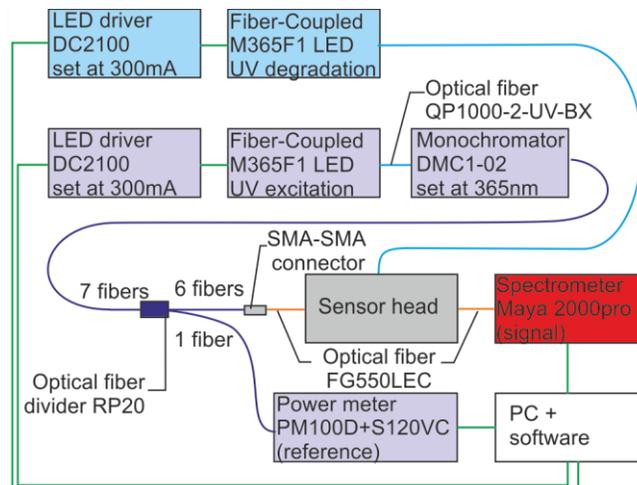


Figure 5. Scheme of the capillary sensor set-up with two UV LEDs.

The utilized monochromator DMC1-02 enables reducing bandwidth to 2nm and fully rejects the long wavelength residual radiation. But this tuning of the M365F1 LED's optical signal to the accuracy required by fluorescence excitation and reading of diesel fuels introduces 1:135 of signal attenuation.

The UV degradation LED working point is set with current stabilization. Therefore, UV power shift can be present and can affect the obtained results. In presented examination, the degradation power was in range from $1293\mu\text{W}$ up to $1300\mu\text{W}$, while the excitation power was in the $8.98\mu\text{W}$ to $9.34\mu\text{W}$ range. While fuel oxidation parameters values as induction time, calculated in percent, varies for different fuel type much more than any possible variation of the mentioned LED power.

III. EXPERIMENTAL RESULTS OF FUEL EXAMINATIONS

The six types of diesel fuel (DF) were under examination including three base clear fuels of premium summer type stored for 4 months in dark container in room conditions (PSC), fresh premium winter fuel (PWC) and fresh eco winter fuel (EWC). These fuels catalog data point similar oxidation stability, but eco fuels are expected to contain more PAHs than premium ones. Therefore, for examination purposes, fuels were modified with additional improver intended to be antioxidant and waxes solvent. This way auxiliary fuels names can be written as: premium summer with additional improver (PSA), premium winter with additional improver (PWA) and eco winter with additional improver (EWA). Mentioned acronyms are used in tables that gather experimental data.

The examination was performed for fuel in ambient temperature equal to $22^\circ\text{C} \pm 1^\circ\text{C}$. In such conditions and used measuring procedure include data processing obtained result of examination accuracy is better than 280 [a.u.] while single measurement accuracy is 50 [a.u.] for full scale range equal 64000 [a.u.].

A. Diesel fuels examination with one UV source working at 265nm

The first set of examination was performed in set-up from Figure 3. The family of characteristics for the set of fuels is presented in Figures 6-11.

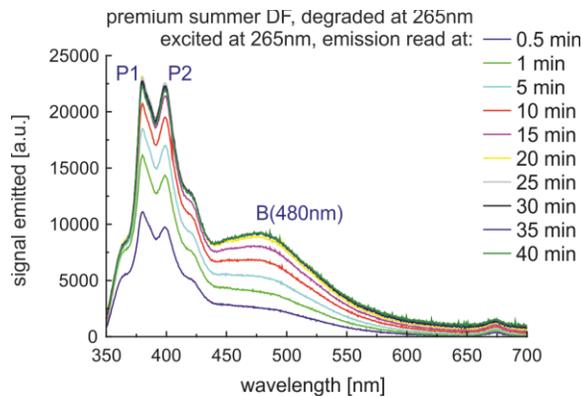


Figure 6. Signal emitted from clear premium summer diesel fuel degraded and excited at 265nm.

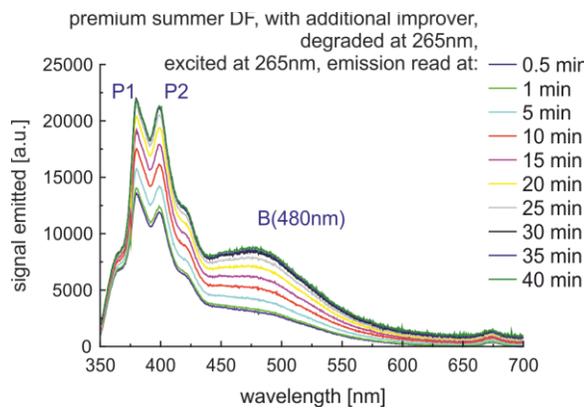


Figure 7. Signal emitted from premium summer diesel fuel with additional improver degraded and excited at 265nm.

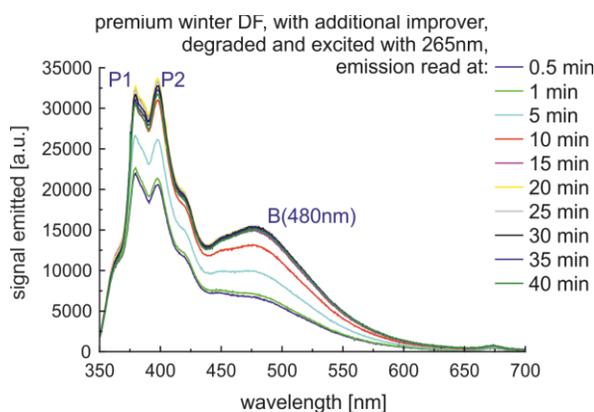


Figure 8. Signal emitted from clear premium winter diesel fuel degraded and excited at 265nm.

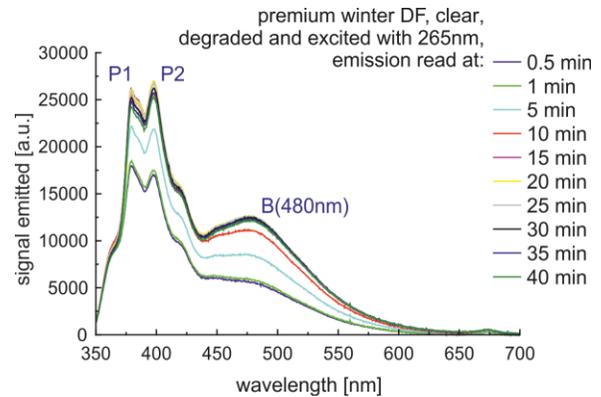


Figure 9. Signal emitted from premium winter diesel fuel with additional improver degraded and excited at 265nm.

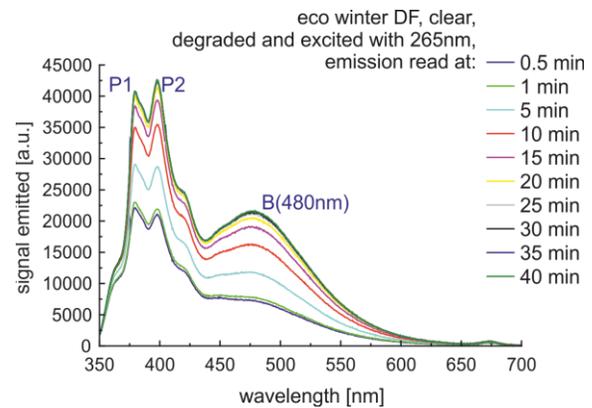


Figure 10. Signal emitted from clear eco winter diesel fuel degraded and excited at 265nm.

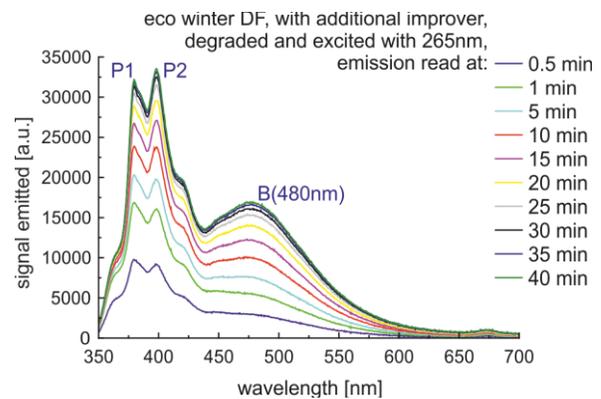


Figure 11. Signal emitted from eco winter diesel fuel with additional improver degraded and excited at 265nm.

The shapes of the characteristics are similar. The characteristic points may be proposed as two peaks P1 at 380nm and P2 at 399nm as well as band signal B480 emitted round 480 nm. For analysis purpose, the characteristics

points P1, P2 and B480 for different fuel types are gathered in Figures 12-14.

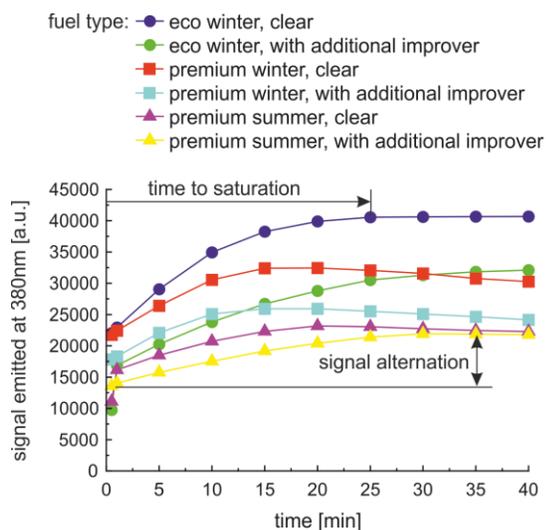


Figure 12. P1 emission at 380nm, excited at 265nm.

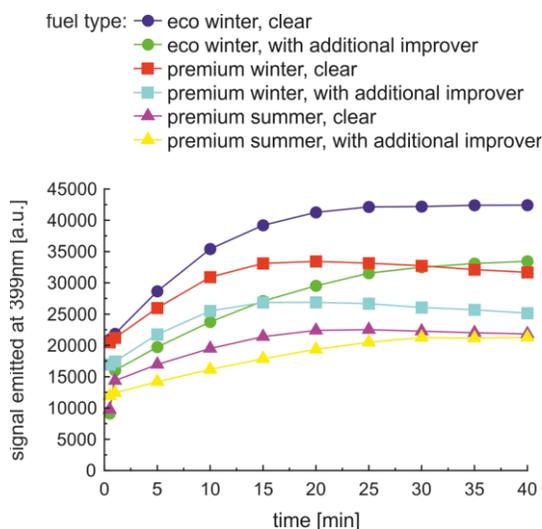


Figure 13. P2 emission at 399nm, excited at 265nm.

The first step of data analysis at Figure 12 should be performed for fuel couples: clear fuel and corresponding fuel with additional improver. In such case the influence of the improver is visible as characteristics shift down with the improver's presence. This shift is not constant, but characteristics pairs are recognizable. The second step of analysis should be performed for fuel types -clear and with additional improver. The clear fuel types characteristics analysis lead to the conclusion that eco fuel generates greater signal at 380nm than premium fuels. But for fuels with improver this statement is not always true.

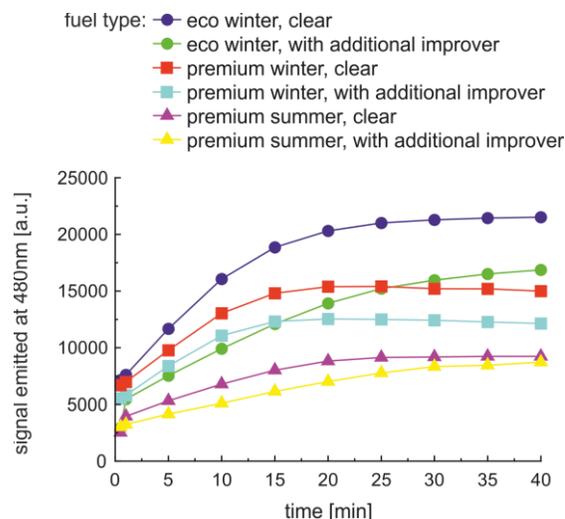


Figure 14. B480 emission at 480nm, excited at 265nm.

The characteristic data points of fuel under with 265nm radiation degradation, counting from the first minute, maximum signal, signal alternation and time to signal saturation at selected wavelength of P1, P2 and B480, are gathered in Table I. The greater is the maximum signal the fuel is expected to contain greater volume of fluorescence particles as PAHs, and therefore be of lower quality. For example, clear eco fuel is described by greater values than the eco fuel with improver, clear premium fuels are characterized by greater values than premium fuels with improver.

TABLE I. DIESEL FUEL EXAMINATION RESULTS WITH ONE UV SOURCE WORKING AT 265NM

Fuel type	Maximum signal during degradation [a.u.]			Signal alteration [a.u.]			Time to saturation [min]		
	P1	P2	B480	P1max -P1min	P2max -P2min	B480max -B480min	P1	P2	B480
EWC	40676	42419	21518	17733	20563	13907	25	35	35
EWA	32101	33456	16870	15236	17455	11412	40	40	40
PWC	32448	33423	15414	10021	12199	8429	15	20	20
PWA	25929	26872	12543	7626	9447	6731	15	20	20
PSC	23181	22510	9242	7031	8148	5294	30	25	30
PSA	21925	21292	8726	7891	8869	5489	40	40	40

The signals alternation during degradation can be used for description of fuel stability. Signals changes for all parameters: P1, P2 and B480 indicate that eco fuel changes are greater than for the premium ones. But for premium fuels the signal change dependency on improver addition is not clear. Despite, the expected relation of lower stability of eco fuel compared to premium is confirmed.

Times to saturation value of P1 increase or stay constant when the improver is added to fuel. Similar dependency has been obtained for values of P2 and B480.

B. Diesel fuels examination with two UV sources working at 365nm

Second set of examination was performed in the set-up from Figure 5. The family of characteristics for the set of fuels is presented in Figures 15-20.

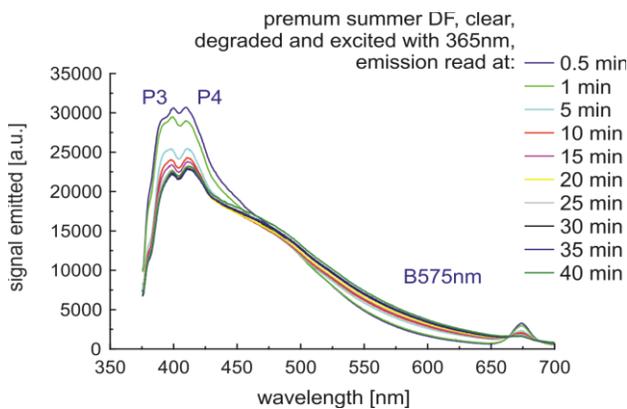


Figure 15. Signal emitted from clear premium summer diesel fuel degraded and excited at 365nm.

The first comparison of characteristics from Figure 15 and Figure 6 shows that peaks are at different wavelength. Therefore the new peaks are numbered as P3 and P4. Moreover, the characteristic changes in B480 bandwidth presented for degradation with 265 nm wavelength now is not observable. Now, the bandwidth of visible changes is from 525nm to 625nm with the greatest variations at about 575nm (B575).

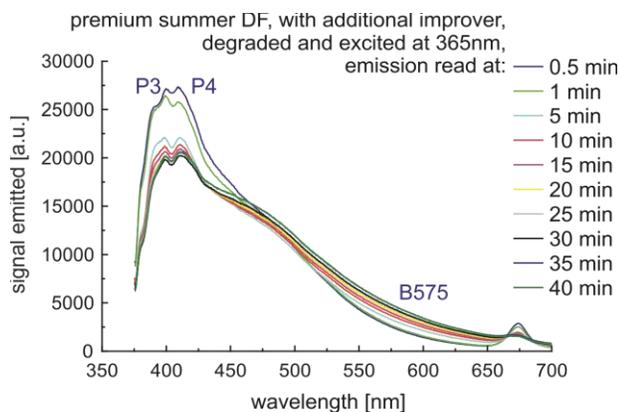


Figure 16. Signal emitted from premium summer diesel fuel with additional improver, degraded and excited at 365nm.

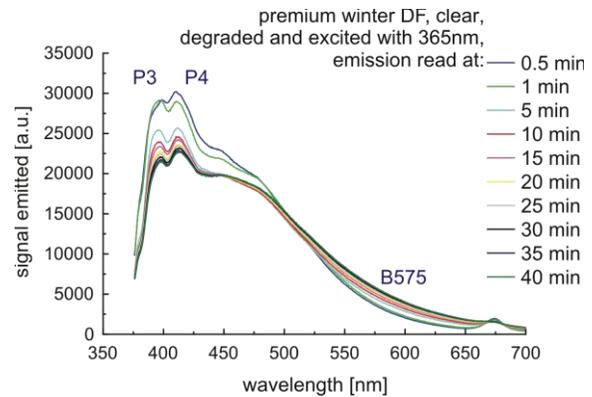


Figure 17. Signal emitted from clear premium winter diesel fuel degraded and excited at 365nm.

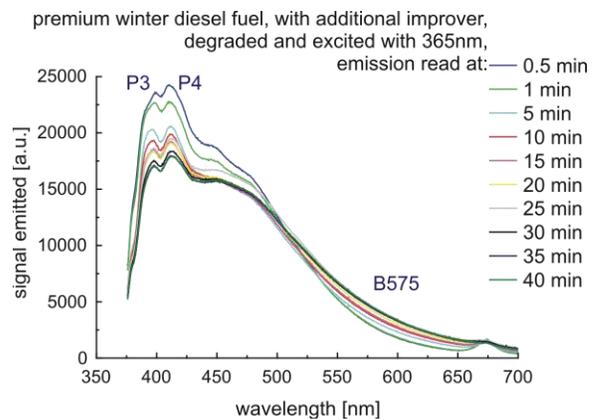


Figure 18. Signal emitted from premium winter diesel fuel with additional improver degraded and excited at 365nm.

Now, the experimental data analysis leads to the conclusion that peak radiation P3 and P4 wavelengths are characterized by minor, but measurable shifts.

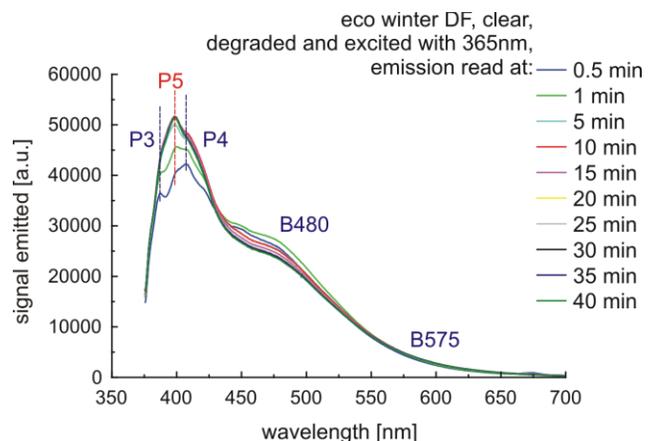


Figure 19. Signal emitted from clear eco winter diesel fuel degraded and excited at 365nm.

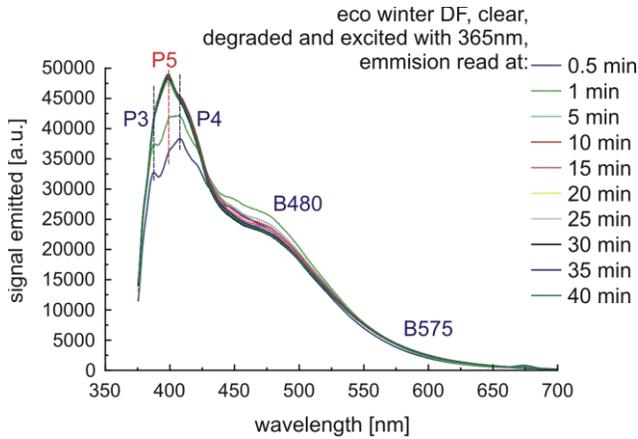


Figure 20. Signal emitted from eco winter diesel fuel with additional improver degraded and excited at 365nm.

The characteristics measured for degradation of eco winter fuel, see Figure 19 and Figure 20, present the new peak (P5) creation as well as P3 and P4 peaks decay.

Besides, the measuring P3 and P4 peaks wavelength shifts seems to be difficult in a sensor device implementation as it requires a spectrometer or a dedicated set of optical filters. But, observation of data gathered in Figures 15-20, leads to the conclusion that amplitude shifts of all mentioned peaks are going in the same direction. Therefore, the wavelength band information (B405) in the range of 380nm to 430nm should illustrate the tendency of signal changes, as presented in Figure 21. In such approximation, the difference between premium and eco type diesel fuel is clear and is the horizontal line at the 30000 [a.u.] of emitted signal.

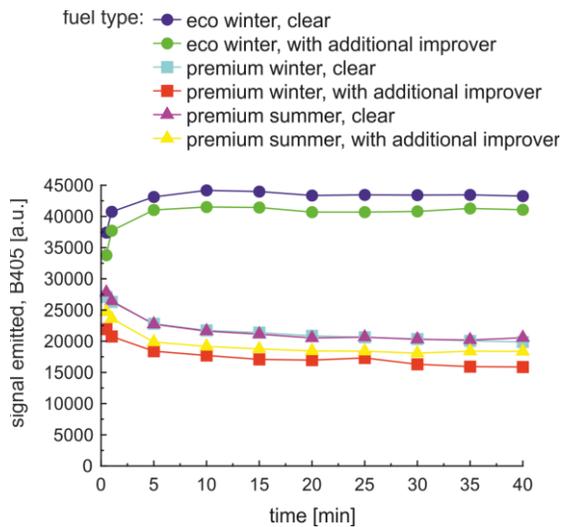


Figure 21. B405 emission in 380nm to 430nm range, excited at 365nm.

The B480 emission for 365 nm excitation seems to be stable, as presented in Figure 22. The most visible changes of signals are for initial period for eco fuel degradation.

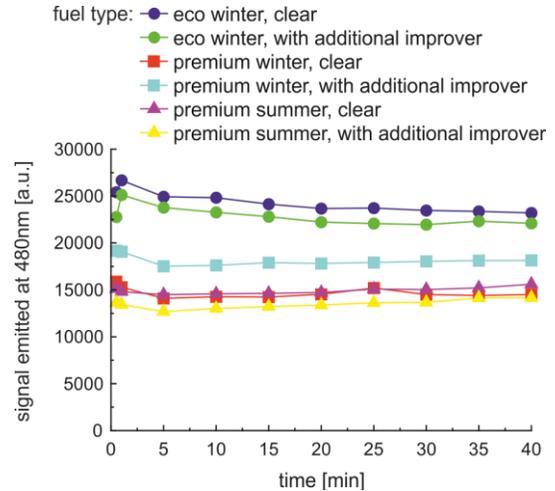


Figure 22. B480 emission excited at 365nm.

Emission at 575 nm, presented in Figure 23, looks different from the characteristic presented in Figure 22. Now, the most stable signals are for eco fuel. But, in both cases the time for saturation may be used as a local fuel stability pointer.

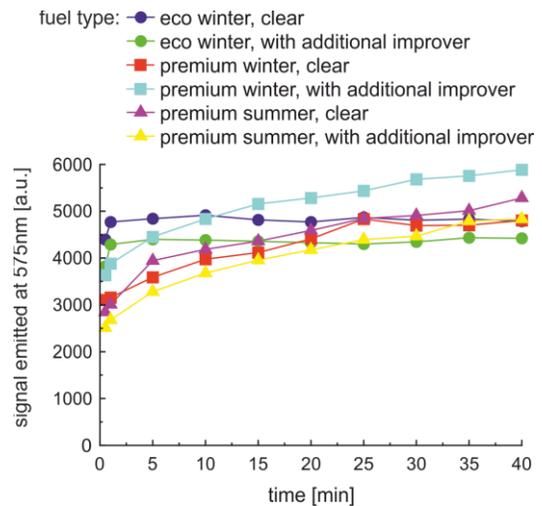


Figure 23. B575 emission excited at 365nm.

The characteristic data points of fuel under with 365 nm radiation degradation, counting from the first minute, are: maximum signal, signal alternation and time to signal saturation at selected wavelength of B405, B480 and B575. They are gathered in Table II.

The most visible aspects here is that the time to saturation that is defined as to the moment when the signal stops to increase, cannot be specified here for some cases. As for example the maximum signal during degradation at 405 nm bandwidth increases only for eco fuel. The maximum signal fuels decreased in 405 nm bandwidth with the use of additional stability improver.

TABLE II. DIESEL FUEL EXAMINATION RESULTS WITH TWO UV SOURCES WORKING AT 365NM

Fuel type	Maximum signal during degradation [a.u.]			Signal alteration [a.u.]			Time to saturation [min]		
	B405	B480	B575	B405max -B405min	B480max -B480min	B575max -B575min	B405	B480	B575
EWC	44178	26681	4914	3438	3473	143	10	-	10
EWA	41501	25118	4435	3780	3160	146	10	-	5
PWC	26308	19080	5885	6429	1549	2003	-	-	25
PWA	20767	15297	4832	4898	1201	1675	-	-	40
PSC	26437	15614	5282	6250	1123	2271	-	-	40
PSA	23616	14160	4832	5544	1478	2152	-	-	40

As previously (see Table I), the peak signals of eco fuels in B405 bandwidth are greater than for premium fuels. The signals changes suggest an opposite classification of chemical stability than expected, because lower values are observed for eco fuels than premium with except of B480 signal alternation. But, signal alternation is relatively small, calculated in percent versus B405 maximum signal.

IV. CONCLUSIONS

The experiments confirm that fuel fluorescence is a sensitive indicator of photochemical transformations. The UV radiation can be used for fuel sample degradation and fluorescence excitation. The sample irradiation at 265 nm and 365 nm results in different emitted signals changes. The 265 nm irradiation results in amplitude shift of two peaks at 380 nm and 399 nm, and also results in signal development at 480 nm. The sample irradiation at 365 nm results sometimes in the transformation of two peaks to one peak centered at 405 nm. The 405 nm characteristic signal emitted during this examination may be direct related with the generation of benzo[a]pyrene.

Comparing the results gathered in Table I and Table II, we can see that signals variations of P1 and P2 peak registered for degradation with 265 nm wavelength are greater than signal variations of B405 observed for degradation with 365 nm. Therefore, chemical stability of fuel is easier to observation for photo degradation with the use of 265 nm wavelength than with the use of 365 nm. The results of data analysis during fuel irradiation at 265 nm enable to claim, that 40 minutes of UV degradation and sequential fluorescence reading at 10 selected moments of time coupled with data processing is enough to determine the chemical stability of the diesel fuel.

Moreover, the ratio of B405 signals of eco fuels versus premium fuels is about 2:1 for excitation with 365 nm wavelength and is semi-stable during degradation. Therefore, B405 signals can be used to pointing PAHs concentration. In this case, the examination does not have to last more than 15 minutes.

ACKNOWLEDGMENT

This work was supported by the EU Horizon 2020 grant of Teaming for Excellence project ID: 664470, "CEZAMAT-Environment" and by the WUT grant no. 8956/E-365/S/2017, task 504/03115/1035/40.000105.

Authors' contribution: Michał Borecki proposed the method of fuel examination, designed the sensor and analyzed the data from heads and set-ups; Mateusz Gęca proposed the algorithm of time-dependent fuels examination and performed the tests of fuels; Przemysław Prus developed the methods of data processing; Michael L. Korwin-Pawlowski counseled the design process, verified the data analysis and the construction specifications; Jan Szmids supervised the working plan.

REFERENCES

- [1] M. Borecki, M. Gęca, M. L. Korwin-Pawlowski, and P. Prus, "Capillary sensor with UV-forced degradation and fluorescence reading of diesel and biodiesel fuel chemical stability," The Eighth International Conference on Sensor Device Technologies and Applications (SENSORDEVICES 2017) IARIA, Aug. 2017, pp. 25-30, ISBN: 978-1-61208-581-4.
- [2] J. Bacha et al., Diesel fuels technical review, Chevron Corporation, San Ramon, CA, USA, 2007.
- [3] S. Jain and M. P. Sharma, "Stability of biodiesel and its blends: a review," Renewable and Sustainable Energy Reviews, vol. 14(2), pp. 667-678, 2010.
- [4] C.I.O. Kamalu, A.C. Anarah, A.O. Ogah, P.C. Okere, J.C. Obijiaku, F.N. Uzundu, "Kinetic modelling of the combustion of aliphatic hydrocarbons," International Journal of Engineering Research and Allied Sciences, vol. 02, pp. 1-15, 2017.
- [5] D. Patara, K.L. Sireesha, A.K. Mishra, "Determination of synchronous fluorescence scan parameters for certain petroleum products," Journal of Scientific & Industrial Research, vol. 59, pp. 300-305, 2000.
- [6] B. L. Smith, L. S. Ott, and Thomas J. Bruno, "Composition-explicit distillation curves of commercial biodiesel fuels:

- comparison of petroleum-derived fuel with B20 and B100," *Ind. Eng. Chem. Res.*, vol. 47, pp. 5832-5840, 2008.
- [7] M. G. Fonseca et al., "Advanced techniques in soybean biodiesel," H. El-Shemy (Ed.), *Soybean - Bio-Active Compounds*, InTech, Rijeka, Croatia, 2013.
- [8] R. A. Korus, T. L. Mousetis, and L. Lloyd, "Polymerization of vegetable oils," *American Society of Agricultural Engineering*, vol. 4, pp. 127-137, 1982.
- [9] Technical Committee of Petroleum Additive Manufactures in Europe, "Fuel additives: use and benefits," ATC Document, vol. 113, pp. 1-68, 2013.
- [10] C.C. Gaylarde, F.M. Bento, and J. Kelly, "Microbial contamination of stored hydrocarbon fuels and its control," *Rev. Microbiol.*, vol. 30, pp. 1-10, 1999.
- [11] L. F. Bautista, C. Vargas, N. González, M. C. Molina, R. Simarro, A. Salmerón, and Y. Murillo, "Assessment of biocides and ultrasound treatment to avoid bacterial growth in diesel fuel," *Fuel Processing Technology*, vol. 152, pp. 56-63, 2016.
- [12] D. E. Nicodem, C. L. B. Guedes, R. J. Correa, and M. C. Z. Fernandes, "Photochemical processes and the environmental impact of petroleum spills," *Biogeochemistry*, vol. 39(2), pp. 121-138, 1997.
- [13] D. L. Plata, C. M. Sharpless, and C. M. Reddy, "Photochemical degradation of polycyclic aromatic hydrocarbons in oil films," *Environmental Science and Technology*, vol. 42(7), pp. 2432-2438, 2008.
- [14] A. Dwivedi, A. Kumar, and J.L. Bhat, "Effect of UV radiation on the growth and petroleum hydrocarbon degradation ability of bacteria," *Oct. Jour. Env. Res.*, vol. 5(1), pp. 032-040, 2017.
- [15] R. Lin and L.L. Tavlarides, "Thermal stability and decomposition of diesel fuel under subcritical and supercritical conditions," *Journal of Supercritical Fluids*, vol. 75, pp. 101-111, 2013.
- [16] J. Czarnocka and M. Odziemkowska, "Diesel fuel degradation during storage process," *Chemik*, vol. 69, pp. 771-776, 2015.
- [17] J.P. Wan, "Properties of fats and oils," in *Introduction to Fat and Oils Technology*, AOCS Press, Champaign Illinois, pp. 20-49, 2000.
- [18] L. Zhang, P. Li, Z. Gong, and X. Li, "Photocatalytic degradation of polycyclic aromatic hydrocarbons on soil surfaces using TiO₂ under UV light," *Journal of Hazardous Materials*, vol. 158, pp. 478-484, 2008.
- [19] O. T. Woo, W. K. Chung, K. H. Wong, A. T. Chow, and P. K. Wong, "Photocatalytic oxidation of polycyclic aromatic hydrocarbons: Intermediates identification and toxicity testing," *Journal of Hazardous Materials*, vol. 168, pp. 1192-1199, 2009.
- [20] M. Nandi and D. Jacobs, "Cetane response of di-tertiary-butyl peroxide in different diesel fuels," *SAE Technical Paper*, vol. 952368, pp. 952368:1-18, 1995.
- [21] H. O. Pritchard and P. Q. E. Clothier, "Anaerobic operation of an internal combustion engine," *J. Chem. Soc. Chem. Commun.*, vol. 20, pp. 1529-1530, 1986.
- [22] 2EHN Industry Work Group, "2-Ethylhexyl nitrate (2EHN) best practices manual," ATC, doc. 79, pp. 1-24, 2004.
- [23] H. Bornemann, F. Scheidt, and W. Sander, "Thermal Decomposition of 2-Ethylhexyl Nitrate (2-EHN)," *Int. J. of Chemical Kinetics*, vol. 34, pp. 34-38, 2002.
- [24] M. Insausti and B. S. Fernández-Band, "Determination of 2-ethylhexyl nitrate in diesel oil using a single excitation emission fluorescence spectra (EEF) and chemometrics analysis," *J. Fundamentals of Renewable Energy and Applications*, vol. 4, pp. 1000137:1-4, 2014.
- [25] C. Schaschke, I. Fletcher, and N. Glen, "Density and Viscosity Measurement of Diesel Fuels at Combined High Pressure and Elevated Temperature," *Processes*, vol. 1, pp. 30-48, 2013.
- [26] G. Fannin, "Distillation process analyser with ASTM D86 Compliance," *Petro Industry News*, no. August/September, pp. 40-41, 2013.
- [27] P. J. García Moreno, R. Perez-Galvez, A. Guadix, and E. M. Guadix, "Influence of the parameters of the Rancimat test on the determination of the oxidative stability index of cod liver oil," *L W T- Food Science and Technology*, vol. 51, pp. 303-308, 2013.
- [28] J. H. Hildebrand, "Factors Determining Chemical Stability," *Chem. Rev.*, vol. 2, pp. 395-417, 1926.
- [29] R. T. Jarviste, R. T. Muoni, J. H. Soone, H. J. Riisalu, and A. L. Zaidentsal, "Diesel fuel oxidation in storage," *Solid Fuel Chemistry*, vol. 42, pp. 123-127, 2008.
- [30] D. Patra and A. K. Mishra, "Study of diesel fuel contamination by excitation emission matrix spectral subtraction fluorescence," *Analytica Chimica Acta*, vol. 454, pp. 209-215, 2002.
- [31] Marilena Meira et al., "Determination of Adulterants in Diesel by Integration of LED Fluorescence Spectra," *J. Braz. Chem. Soc.*, vol. 26, pp. 1351-1356, 2015.
- [32] M. Włodarski, A. Bombalska, M. Mularczyk-Oliwa, M. Kaliszewski, and K. Koczyński, "Fluorimetric techniques in analysis and classification of fuels," *Proc. SPIE 8703*, pp. 87030B:1-7, 2013, DOI: 10.1117/12.2011706.
- [33] M. Rasouli, S. H. Tavassoli, S. J. Mousavi, and S. M. R. Darbani, "Measuring of naphthalene fluorescence emission in the water with nanosecond time delay laser induced fluorescence spectroscopy method," *Optik*, vol. 127, pp. 6218-6223, 2016.
- [34] S. Chikh, Y.B. Pemanos, T. Zourane, and B. Hamada, "Synthesis of a biobased antioxidant additive for diesel fuel," *American Journal of Applied Chemistry*, vol. 2(1), pp. 10-14, 2014.
- [35] C. L. B. Guedes, E. DiMauro, A. De Campos, L. F. Mazzochin, G. M. Bragagnolo, F. A. De Melo, and M. T. Piccinato, "EPR and fluorescence spectroscopy in the photodegradation study of arabian and colombian crude oils," *International Journal of Photoenergy*, vol. 2006, pp. 48462:1-6, 2006, DOI 10.1155/IJP/2006/48462.
- [36] M. Borecki et al., "Fiber optic capillary sensor with smart optode for rapid testing of the quality of diesel and biodiesel fuel," *IJASM*, vol. 7, pp. 57-67, 2014.
- [37] M. Borecki et al., "Dynamical capillary rise photonic sensor for testing of diesel and biodiesel fuel," *Sensors & Transducers Journal*, vol. 193, pp. 11-22, 2015.
- [38] M. Borecki and M.L. Korwin-Pawlowski, "Capillary sensor with UV-VIS reading of effects of diesel and biodiesel fuel degradation in storage," *Sensors & Transducers*, vol. 205, pp. 1-9, 2016.
- [39] A. Baranowska, P. Miluski, M. Kochanowicz, J. Zmojda, and D. Dorosz, "Capillary optical fibre with Sm³⁺ doped ribbon core," *Proc. SPIE*, vol. 9662, pp. 96620Z, 2015.
- [40] M. Geca, M. Borecki, M. L. Korwin-Pawlowski, and A. Kociubiński, "Local liquid sample heating: integration and isolation of a micro-heater," *Proc. SPIE*, vol. 9662, pp. 96620E, 2015.
- [41] A. Kociubiński, M. Borecki, M. Duk, M. Sochacki, and M. L. Korwin-Pawlowski, "3D photodetecting structure with adjustable sensitivity ratio in UV-VIS range," *Microelectronic Engineering*, vol. 154, pp. 48-52, 2016.
- [42] P. Doroz, M. Duk, M. L. Korwin-Pawlowski, and M. Borecki, "Amplifiers dedicated for large area SiC photodiodes," *Proc SPIE*, vol. 10031, pp. 100311R, 2016.

- [43] P. Prus, M. Borecki, M. L. Korwin-Pawłowski, A. Kociubiński, and M. Duk, "Automatic detection of characteristic points and form of optical signals in multiparametric capillary sensors," *Proc. SPIE*, vol. 9290, pp. 929009, 2014.
- [44] M. T. Ibrahim, R. Telford, P. Dini, P. Lorenz, N. Vidovic, and R. Anthony, "Self-adaptability and man-in-the-loop: a dilemma in autonomic computing systems," *Proc. IEEE Computer Society 15th International Workshop on Database and Expert Systems Applications*, Los Alamitos, California, USA, pp. 722-729, 2004.
- [45] V. Zamora, Z. Zhang and A. Meldrum, "Refractometric Sensing of Heavy Oils in Fluorescent Core Microcapillaries," *Oil & Gas Science and Technology – Rev. IFP Energies nouvelles*, vol. 70(3), pp. 487-495, 2015.
- [46] S. Y. Yurish, N. V. Kirianaki, and R. Pallàs - Areny, "Universal frequency - to - digital converter for quasi - digital and smart sensors: specifications and applications," *Sensor Review*, vol. 25(2), pp. 92-99, 2005.
- [47] Z. Bielecki, "Readout electronics for optical detectors," *Optoelectronics Review*, vol. 12, pp. 129-137, 2004.
- [48] P. Firek, A. Werbowy, and J. Szmids, "MIS field effect transistor with barium titanate thin film as a gate insulator," *Materials Science and Engineering B: Solid-State Materials for Advanced Technology*, vol. 165(1-2), pp. 126-128, 2009.
- [49] R. Srinivasan, "Photochemistry of conjugated dienes and trienes," in *Advances in Photochemistry*, vol. 4, Eds. W. A. Noyes, G. S. Hammond and J. N. Pitts, John Wiley & Sons, Hoboken, NJ, USA, 1966.
- [50] C.M. Carruco Laranjeira et al., "Used food oils: physical-chemical indicators of quality degradation," *Proc. FOODBALT 2017*, pp. 154-159, 2017. DOI: 10.22616/foodbalt.2017.040.
- [51] A. M. Rivera-Figueroa, K. A. Ramazan, and B. J. Finlayson-Pitts, "Fluorescence, absorption, and excitation spectra of polycyclic aromatic hydrocarbons as a tool for quantitative analysis," *J. of Chemical Education*, vol. 81, pp. 242-245, 2004.
- [52] C. Brussol, M. Duane, P. Carlier, and D. Kotzias, "Photo-induced OH reactions of naphthalene and its oxidation products on SiO₂," *Environ. Sci. Pollut. Res. Int.*, vol. 6(3), pp. 138-140, 1999.
- [53] S. Hachiya, K. Asai, and G.-i. Konishi, "Unique solvent-dependent fluorescence of nitro-group-containing naphthalene derivatives with weak donor-strong acceptor system," *Tetrahedron Letters*, vol. 54, pp. 1839-1841, 2013.
- [54] L. H. Zhang, C.B. Xu, Z. L. C, and X. M. Li, "Effect of UV Intensity and Wavelength on Photocatalytic Degradation of Polycyclic Aromatic Hydrocarbons," *Advanced Materials Research*, vols. 233-235, pp. 466-469, 2011.
- [55] F. Valerio and A. Lazzarotto, "Photochemical Degradation of Polycyclic Aromatic Hydrocarbons (PAH) in Real and Laboratory Conditions," *International Journal of Environmental Analytical Chemistry*, vol. 23, pp. 135-151, 1985.
- [56] S. Bertilsson and A. Widenfalk, "Photochemical degradation of PAHs in freshwaters and their impact on bacterial growth – influence of water chemistry," *Hydrobiologia* vol. 469, pp. 23–32, 2002.
- [57] D. E. Nicodem, C. L. B. Guedes, and R. J. Correa, "Photochemistry of petroleum I. Systematic study of a Brazilian intermediate crude oil," *Marine Chemistry*, vol. 63, pp. 93-104, 1998.
- [58] I. Beriman, "Handbook of florescence spectra of Aromatic Molecules –2nd Edition" Academic Press, New York, 1971.

Investigating Different Ballistic Threats to Validate the Simulation Model of Fiber-Reinforced Plastics

Experimental Model Validation of Fiber Composites Under Ballistic Impact

Arash Ramezani and Hendrik Rothe

Chair of Measurement and Information Technology

University of the Federal Armed Forces

Hamburg, Germany

Email: ramezani@hsu-hh.de, rothe@hsu-hh.de

Abstract—In the security sector, the partly insufficient safety of people and equipment due to failure of industrial components is an ongoing problem that causes great concern. Since computers and software have spread into all fields of industry, extensive efforts are currently made to improve the safety by applying certain numerical solutions. A fibre-reinforced composite is a promising material for ballistic protection due to its high strength, stiffness and low density. The use of ultra-high molecular weight polyethylene (UHMW-PE) composite as part of the personal armour system has the potential to provide significant weight savings or improved protection levels over traditional metallic materials. Although already used in different applications, both as spall liners and within complex multi-element/multi-material packages, there is a limited understanding of the mechanisms driving ballistic performance. Existing analysis tools do not allow a good approximation of performance, while existing numerical models are either incapable of accurately capturing the response of thick UHMW-PE composite to ballistic impact or are unsuited to model thick targets. In response, this paper aims to identify the key penetration and failure mechanisms of thick UHMW-PE composites under ballistic impact and develop analytical and numerical models that capture these mechanisms and allow accurate prediction of ballistic performance to optimize modern armour systems. An analysis methodology is proposed to model the behaviour of thick UHMW-PE composite panels under ballistic impact using inhomogeneities on the macroscale. A sub-laminate approach for discretisation of the target is proposed to overcome the problems of premature through-thickness failure in the material model. The methodology was extensively validated against existing experimental ballistic impact data and results for UHMW-PE targets. Finally, a numerical modelling methodology was developed for the analysis of thick UHMW-PE composite under ballistic impact.

Keywords—generally valid simulation models; hydrocode analysis; fiber-reinforced plastics; optimization; armor systems; ballistic trials.

I. INTRODUCTION

For thousands of years, natural materials had formed the basis of human existence: clothing, tools, and articles of consumption, all were made from leather, metal, stone, clay, or other substances obtained directly from nature. In

contrast, most of the manmade materials such as porcelain, glass, and metal alloys were discovered more or less by accident. At the beginning of the twentieth century, dwindling deposits of important resources and their escalating prices triggered off an intensive search for synthetic, or manmade, substitute materials. The demand from the fast-growing industries was increasing in line with fundamental technical changes and could no longer be satisfied with natural materials alone. In time, countless compounds, including a high number of plastics, were synthesized from naturally occurring raw materials such as coal, coal tar, crude oil, and natural gas.

The object behind combining different materials to form a composite with enhanced properties and synergetic effects is par for the course in nature. A section through a paracortical cell in merino wool or through a bamboo stems exhibits structures similar to the micrograph of a unidirectional carbon-fibre-reinforced epoxy resin (CF-EP). Not only in the microstructure can nature be seen as the progenitor of fibre-reinforced plastics, but also in the application of lightweight design principles.

Why material scientists integrate fibers in materials to such advantage can be answered by the following four paradoxes of engineering materials:

- The paradox of the solid material: The actual strength of a solid material is very much lower than the calculated theoretical value.
- The paradox of the fiber form: The strength of a material in fiber form is many times higher than that of the same material in another form, and the thinner the fiber, the greater the strength.
- The paradox of the free clamped length: The shorter the length between the clamps, the greater the strength measured on the test piece.
- The paradox of composites: When taken as a whole, a composite can withstand stresses that would fracture the weaker component, whereas the composite's stronger component can exhibit a greater percentage of its theoretical strength than when loaded singly.

So, the principle of combining different materials to form a composite with enhanced properties is just as

common in nature as it is in lightweight engineering. This design method based on nature's example has virtually revolutionized many fields of technology, with the result that they can now utilize the superior properties of high-tensile, lightweight materials for the first time.

This work will focus on fiber-reinforced plastics, more precisely composite armor structures consisting of several layers of ultra-high molecular weight polyethylene (UHMW-PE), a promising ballistic armor material due to its high specific strength and stiffness. First approaches are discussed in detail in [1].

UHMW-PE is a thermoplastic polymer made from very long molecular chains of polyethylene. Thermoplastics soften when subjected to heat and so can be repeatedly remoulded. Cut-offs can be remelted and introduced back into the production process. Many thermoplastics are soluble in organic solvents. Thermoplastics can be joined by welding under the application of heat or by the action of solvents [2].

Figure 1 shows the chemical structure of polyethylene, where in UHMW-PE the number of repeated chains (n) is in the order of 10^5 , giving rise to molecular weights in the order of 10^6 [3]. As a non-polar molecule, interaction between polyethylene molecules is given by very weak Van der Waals forces.

However, due to the ultra-long polymer chain, significant strength can be derived through a gel spinning process that produces highly oriented and crystalline molecular structures aligned in the spinning direction. The gel spinning process firstly involves dissolving UHMW-PE in a solvent at high temperature. The solution is then pushed through a spinneret to form liquid filament that is then quenched in water to form gel-fibers. These fibers are then drawn in hot air at high strain rates of the order of 1 s^{-1} forming fibers with smooth circular cross-sections approximately 17 mm in diameter [4] with a molecular orientation greater than 95% and a crystallinity of up to 85% [3], see Figure 2.

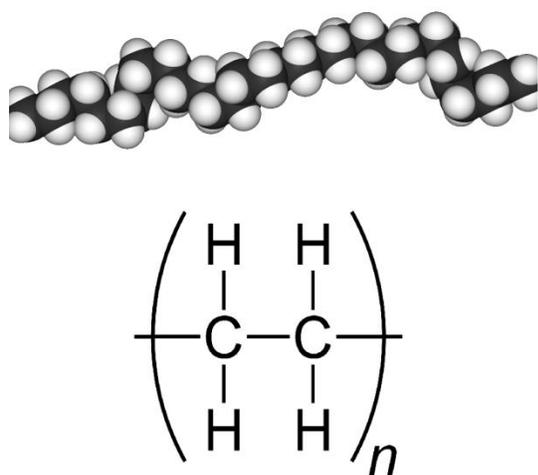


Figure 1. Skeletal formula and spacefill model of a polyethylene.

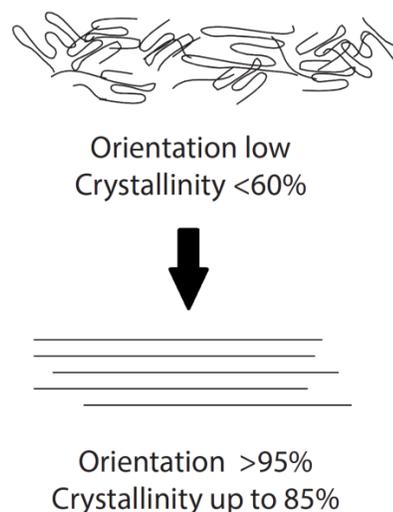


Figure 2. Increase in molecular orientation and crystallinity through gel spinning of UHMW-PE [5].

These fibers are composed of smaller macro-fibrils approximately 0.5 mm to 2 mm in diameter, which in turn are made of micro-fibrils, 20 nm in diameter. Commercial UHMW-PE fiber is manufactured by, amongst others, Dutch State Mines (DSM) and Honeywell under the trade names Dyneema® and Spectra®, respectively. The fibers are used in a variety of applications requiring high specific strength and low weight. This includes high strength ropes and nets, cut-resistant gloves, as well as blast and ballistic protection.

For ballistic protection applications, the fibers can be woven into fabrics to provide a soft and flexible material or coated in a matrix and aligned to form uni-directional plies, which are then stacked and pressed under temperature and pressure to form rigid laminates.

UHMW-PE composites and fabrics have been shown to be extremely effective against ballistic threats, particularly in weight-critical applications, e.g., personal protection vests and helmets for protection against small calibre threats [6]. The goal is to evaluate the ballistic efficiency of UHMW-PE composite with numerical simulations, promoting an effective development process.

Due to the fact that all engineering simulation is based on geometry to represent the design, the target and all its components are simulated as computer-aided design (CAD) models. The work will also provide a brief overview of ballistic tests to offer some basic knowledge of the subject, serving as a basis for the comparison of the simulation results. Details of ballistic trials on composite armor systems are presented. Instead of running expensive trials, numerical simulations should identify vulnerabilities of structures. Contrary to the experimental result, numerical methods allow easy and comprehensive studying of all mechanical parameters. Modelling will also help to understand how the armor schemes behave during impact

and how the failure processes can be controlled to our advantage. By progressively changing the composition of several layers and the material thickness, the composite armor will be optimized.

There is every reason to expect possible weight savings and a significant increase in protection, through the use of numerical techniques combined with a small number of physical experiments.

This work deals with numerical simulations of impact problems on fiber-based composite armor using the commercial finite-element-code ANSYS AUTODYN. Having presented some basic knowledge on the theory of numerical simulation in AUTODYN, two recently published approaches for modelling impact on the selected composite (Dyneema® HB26) are explained. While both of them make use of a nonlinear-orthotropic material model implemented in the AUTODYN-code, they differ in the way how the highly inhomogeneous microstructure of HB26 is represented geometrically. The first approach chooses a fully homogeneous description, whereas the other approach discretizes the composite into sublaminates, which are kinematically joined at the surfaces and breakable when a certain contact-stress is reached. They will be discussed in detail in Section III. In order to validate the two approaches, the response of HB26-samples impacted by handgun-projectiles was determined experimentally and compared to the corresponding numerical results. Unfortunately, a poor agreement between experimental and numerical results was found, which gave rise to the development of an alternative modelling approach. In doing so, the composite was subdivided into alternating layers of two different types. While the first type of layers was modeled with open-literature properties of UHMW-PE-fibers, polymer-matrix-behavior was assigned to the second type. Having adjusted some of the parameters, good agreement between experiment and simulation was found with respect to residual velocity and depth of penetration for the considered impact situations.

After a brief introduction and description of the principles of simulation in Section II, state-of-the-art models of fiber-reinforced plastics are discussed in Section III. There is a short section on ballistic trials where the experimental set-up is depicted, followed by Section V describing the model validation. Section VI presents the analysis with numerical simulations and the results of this work. The paper ends with a concluding paragraph in Section VII and an outlook in Section VIII.

II. PRINCIPLES OF SIMULATION

To deal with problems involving the release of a large amount of energy over a very short period of time, e.g., explosions and impacts, there are three approaches: as the problems are highly non-linear and require information regarding material behavior at ultra-high loading rates which is generally not available, most of the work is *experimental* and thus may cause tremendous expenses.

Analytical approaches are possible if the geometries involved are relatively simple and if the loading can be described through boundary conditions, initial conditions or a combination of the two. *Numerical* solutions are far more general in scope and remove any difficulties associated with geometry [7]. They apply an explicit method and use very small time steps for stable results.

For problems of dynamic fluid-structure interaction and impact, there typically is no single best numerical method which is applicable to all parts of a problem. Techniques to couple types of numerical solvers in a single simulation can allow the use of the most appropriate solver for each domain of the problem.

The goal of this paper is to evaluate a hydrocode, a computational tool for modelling the behavior of continuous media. In its purest sense, a hydrocode is a computer code for modelling fluid flow at all speeds [8]. For that reason, a structure will be split into a number of small elements. The elements are connected through their nodes (see Figure 3). The behavior (deflection) of the simple elements is well-known and may be calculated and analyzed using simple equations called shape functions [9].

By applying coupling conditions between the elements at their nodes, the overall stiffness of the structure may be built up and the deflection/distortion of any node – and subsequently of the whole structure – can be calculated approximately [10].

A repeatedly used term in this context will be discretization. Its meaning is that equations, formulated to continuously describe a function or functional in space and time, is solved only at certain discrete locations and instants of time [11]. The most commonly used spatial discretization methods are Lagrange, Euler, ALE (a mixture of Lagrange and Euler), and mesh-free methods, such as Smooth Particles Hydrodynamics (SPH) [12].

A. Lagrange

The Lagrange method of space discretization uses a mesh that moves and distorts with the material it models as a result of forces from neighboring elements (meshes are imbedded in material). There is no grid required for the external space, as the conservation of mass is automatically satisfied and material boundaries are clearly defined. This is the most efficient solution methodology with an accurate pressure history definition.

The Lagrange method is most appropriate for representing solids, such as structures and projectiles. If however, there is too much deformation of any element, it results in a very slowly advancing solution and is usually terminated because the smallest dimension of an element results in a time step that is below the threshold level.

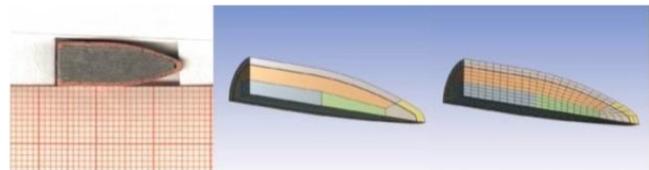


Figure 3. Example grid.

B. Euler

The Euler (multi-material) solver utilizes a fixed mesh, allowing materials to flow (advection) from one element to the next (meshes are fixed in space). Therefore, an external space needs to be modeled. Due to the fixed grid, the Euler method avoids problems of mesh distortion and tangling that are prevalent in Lagrange simulations with large flows. The Euler solver is very well-suited for problems involving extreme material movement, such as fluids and gases. To describe solid behavior, additional calculations are required to transport the solid stress tensor and the history of the material through the grid. Euler is generally more computationally intensive than Lagrange and requires a higher resolution (smaller elements) to accurately capture sharp pressure peaks that often occur with shock waves.

C. ALE

The ALE method of space discretization is a hybrid of the Lagrange and Euler methods. It allows redefining the grid continuously in arbitrary and predefined ways as the calculation proceeds, which effectively provides a continuous rezoning facility. Various predefined grid motions can be specified, such as free (Lagrange), fixed (Euler), equipotential, equal spacing, and others. The ALE method can model solids as well as liquids. The advantage of ALE is the ability to reduce and sometimes eliminate difficulties caused by severe mesh distortions encountered by the Lagrange method, thus allowing a calculation to continue efficiently. However, compared to Lagrange, an additional computational step of rezoning is employed to move the grid and remap the solution onto a new grid [13].

D. SPH

The mesh-free Lagrangian method of space discretization (or SPH method) is a particle-based solver and was initially used in astrophysics. The particles are imbedded in material and they are not only interacting mass points but also interpolation points used to calculate the value of physical variables based on the data from neighboring SPH particles, scaled by a weighting function. Because there is no grid defined, distortion and tangling problems are avoided as well. Compared to the Euler method, material boundaries and interfaces in the SPH are rather well defined and material separation is naturally handled. Therefore, the SPH solver is ideally suited for certain types of problems with extensive material damage and separation, such as cracking. This type of response often occurs with brittle materials and hypervelocity impacts. However, mesh-free methods, such as Smooth Particles Hydrodynamics, can be less efficient than mesh-based Lagrangian methods with comparable resolution.

ANSYS Autodyn lets you select from these different solver technologies so the most effective solver can be used for a given part of the model. Figure 4 gives a short overview of the solver technologies mentioned above. The crucial factor is the grid that causes different outcomes.

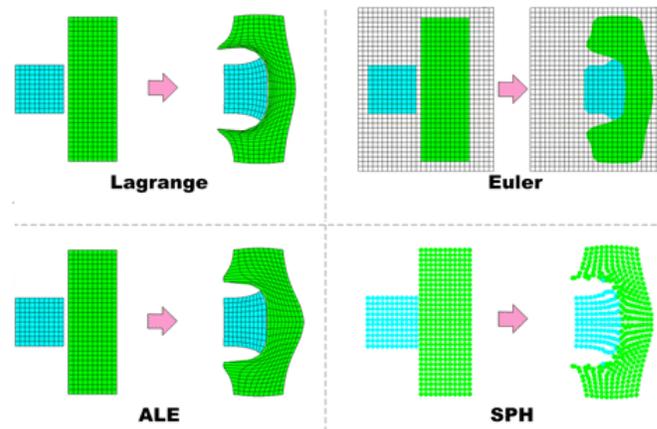


Figure 4. Examples of Lagrange, Euler, ALE, and SPH simulations on an impact problem [14].

Using a CAD-neutral environment that supports bidirectional, direct, and associative interfaces with CAD systems, the geometry can be optimized successively [17]. Therefore, several runs are necessary: from modelling to calculation to the evaluation and subsequent improvement of the model (see Figure 5).

AUTODYN's interaction logic enables automatic communication between the various solvers coexisting within the same model. Lagrange-Lagrange, SPH-Lagrange and Euler-Lagrange interactions can all be created within the model in a simple and intuitive manner. This allows fluid structure interactions to be simulated. Furthermore, it can be combined with AUTODYN's extensive remapping and de-zoning capabilities between the different solvers and a wide range of erosion settings. It is also possible to retain inertia of eroded material. In this case, the mass and momentum of the free node is retained and can be involved in subsequent impact events to transfer momentum in the system.

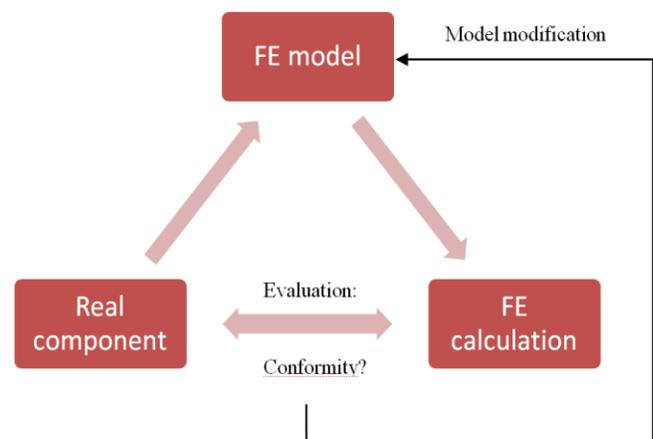


Figure 5. Basically iterative procedure of a FE analysis [15].

III. STATE-OF-THE-ART

The numerical modelling of composite materials under impact can be performed at a constituent level (i.e., explicit modelling of fiber and matrix elements, e.g., [9]), a meso-mechanical level (i.e., consolidated plies or fiber bundles, e.g., [10]), or macromechanically in which the composite laminate is represented as a continuum.

In [11–14] a non-linear orthotropic continuum material model was developed and implemented in a commercial hydrocode (i.e., ANSYS AUTODYN) for application with aramid and carbon fiber composites under hypervelocity impact. The non-linear orthotropic material model includes orthotropic coupling of the material volumetric and deviatoric responses, a non-linear equation of state (EoS), orthotropic hardening, combined stress failure criteria and orthotropic energy-based softening. For more detail refer to [15]. Lässig et al. [16] conducted extensive experimental characterization of Dyneema® HB26 UHMW-PE composite for application in the continuum non-linear orthotropic material model, and validated the derived material parameters through simulation of spherical projectile impacts at hypervelocity.

A number of researchers have applied the non-linear orthotropic model for UHMW-PE composites with varying levels of success (Hayhurst et al. [17], Herlaar et al. [18], Ong et al. [19], Heisserer and Van der Werff [20] and Lässig et al. [16]). Ong et al. [19] assumed material properties of UHMW-PE composite based on those of Kevlar® with some data from literature, which resulted in poor predictions of the penetration behaviour. Hayhurst et al. [17], Herlaar et al. [18] and Heisserer and Van der Werff [20] used material input parameters derived from a range of experiments, and reported better prediction, although the results cannot be independently verified because the material parameters are not provided.

Nguyen et al. [21] evaluated and refined the modelling approach and material model parameter set developed in [16] for the simulation of impact events from 400 m/s to 6600 m/s. Across this velocity range the sensitivity of the numerical output is driven by different aspects of the material model, e.g., the strength model in the ballistic regime and the equation of state (EoS) in the hypervelocity regime.

This paper will present an optimized solution of this problem with an enhanced model for ultra-high molecular weight polyethylene under impact loading. For the first time, composite armor structures consisting of several layers of fiber-reinforced plastics are simulated for all current military threats.

IV. BALLISTIC TRIALS

Ballistics is an essential component for the evaluation of our results. Here, terminal ballistics is the most important sub-field. It describes the interaction of a projectile with its target. Terminal ballistics is relevant for both small and

large caliber projectiles. The task is to analyze and evaluate the impact and its various modes of action. This will provide information on the effect of the projectile and the extinction risk.

Historically, impact events are classified according to the impact velocity. One such approach was proposed by Zukas et al. [22] who categorised impact problems based on impact velocity, where the material response and established strain rate characterised the impact problem. Figure 6 depicts the strain rate, the impact velocity required to achieve the strain rate and the material effects as proposed by Zukas et al. [22].

Under this classical approach, the ballistic regime could be considered to be within the strain rate range of 10^2 s^{-1} to 10^4 s^{-1} , which corresponds to an impact velocity of between 50 m/s to 3000 m/s. Beyond 4000 m/s (depending on the materials), an impact will lead to a complete break-up and melting of the projectile. According to Zukas et al. [22], within this region the material strength is important in resisting penetration, and there is an onset of hydrodynamic effects. This approach to define the impact regime using the strain rate is insufficient when considering the impact behaviour for a diverse range of materials with different properties. For example, very low velocity impact of a hard object into a fluid can be described entirely using hydrodynamic theory, which under Figure 6 would be classified as both a low strain rate (low velocity) and high strain rate (hydrodynamic effects) problem.

On the other hand, Wilbeck (in [23]) proposed classification by the ratio of the impact pressure (P) to the material strength (S).

$$P/S_p \text{ and } P/S_t \tag{1}$$

where subscripts p is the projectile and t is the target. The pressure, P, according to hydrodynamic theory is given by:

$$P \cong \rho V^2 \tag{2}$$

where ρ is material density and V is the impact velocity. From this, impact events can be classified in nine different regimes, depicted in matrix form in Figure 7.

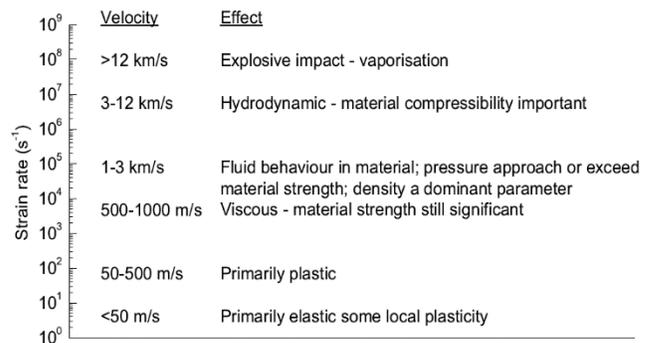


Figure 6. Impact response of high strength materials [22].

	$\frac{P}{S_t} \ll 1$	$\frac{P}{S_t} \sim 1$	$\frac{P}{S_t} \gg 1$
$\frac{P}{S_p} \ll 1$	1	2	3
$\frac{P}{S_p} \sim 1$	4	5	6
$\frac{P}{S_p} \gg 1$	7	8	9

Figure 7. Matrix of impact regime [23]. The ballistic regimes are shaded.

The ballistic regime is considered to reside in regime 2, 4, 5, 6, and 8 according to Figure 7, where the impact pressure is close to the material strength of the target, projectile or both. For this paper, regime 5, 6 and 8 are of interest because at impact velocities typical of FSPs and projectiles, the impact pressures are considered to be on the order of, or greater than the strength of UHMW-PE composite.

Given that a projectile strikes a target, compressive waves propagate into both the projectile and the target. Relief waves propagate inward from the lateral free surfaces of the penetrator, cross at the centerline, and generate a high tensile stress. If the impact was normal, we would have a two-dimensional stress state. If the impact was oblique, bending stresses will be generated in the penetrator. When the compressive wave reached the free surface of the target, it would rebound as a tensile wave. The target may fracture at this point. The projectile may change direction if it perforates (usually towards the normal of the target surface).

Because of the differences in target behavior based on the proximity of the distal surface, we must categorize targets into four broad groups. A semi-infinite target is one where there is no influence of distal boundary on penetration. A thick target is one in which the boundary influences penetration after the projectile is some distance into the target. An intermediate thickness target is a target where the boundaries exert influence throughout the impact. Finally, a thin target is one in which stress or deformation gradients are negligible throughout the thickness.

There are several methods by which a target will fail when subjected to an impact. The major variables are the target and penetrator material properties, the impact velocity, the projectile shape (especially the ogive), the geometry of the target supporting structure, and the dimensions of the projectile and target.

In order to develop a numerical model, a ballistic test program is necessary. The ballistic trials are thoroughly documented and analyzed – even fragments must be collected. They provide information about the used armor and the projectile behavior after fire, which must be consistent with the simulation results.

In order to create a data set for the numerical simulations, several experiments have to be performed. Ballistic tests are recorded with high-speed videos and analyzed afterwards. Testing was undertaken at an indoor ballistic testing facility. The target stand provides support behind the target on all four sides. Every ballistic test program includes several trials with different composites. The set-up has to remain unchanged.

The camera system is a PHANTOM v1611 that provides the versatility and flexibility needed in a variety of applications. With the proprietary widescreen CMOS sensor, the v1611 can acquire and save up to 16 gigapixels-per-second of data. That means at its full megapixel resolution of 1280×800, it can achieve 16,000 frames-per-second (fps). At reduced resolutions, the v1611 offers frame rates of 646,000 fps. With an internal mechanical shutter, the black frame can be obtained by simply closing the shutter. No physical access to the camera is needed. Ease-of-Use features include common signal connections which are conveniently located on its back panel, and include connections for timecode, dual power inputs, HD-SDI, a GPS input, frame synchronization, and trigger.

The use of a polarizer and a neutral density filter is advisable, so that waves of some polarizations can be blocked while the light of a specific polarization can be passed.

Several targets of different laminate configurations were tested to assess the ballistic limit (V_{50}). The ballistic limit is considered the velocity required for a particular projectile to reliably (at least 50% of the time) penetrate a particular piece of material [24]. After the impact, the projectile is examined regarding any kind of change it might have undergone.

V. MODEL VALIDATION

Experimental characterisation of the ballistic performance of UHMW-PE composite can be prohibitively expensive, so it is highly desirable to establish computationally efficient numerical models that accurately predict the ballistic response of the material. First, existing models should be validated.

A. Resources

In [16], numerical simulations of 15 kg/m² Dyneema® HB26 panels impacted by 6 mm diameter aluminum spheres between 2052 m/s to 6591 m/s were shown to provide very good agreement with experimental measurements of the panel ballistic limit and residual velocities, see Figure 8. The modelling approach and material parameter set from [16] were applied to simulate impact experiments at velocities in the ballistic regime (here considered as < 1000 m/s). Lambert-Jonas parameters (a , p , V_{bl}) are provided in the legend.

In Figure 8, the results of modelling impact of 20 mm fragment simulating projectiles (FSPs) against 10 mm thick Dyneema® HB26 are shown. The model shows a significant under prediction of the ballistic limit, 236 m/s compared to 394 m/s.

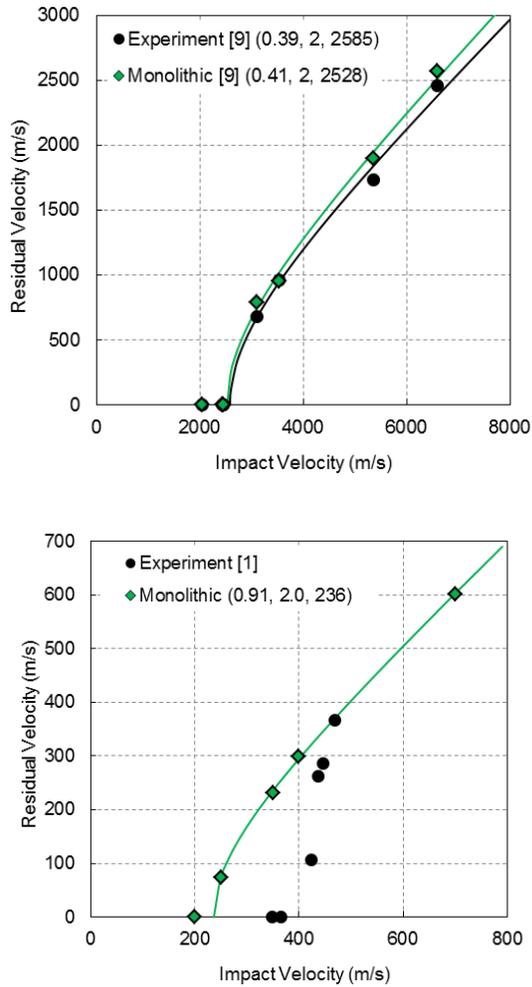


Figure 8. Experimental and numerical impact residual velocity results for impact of 6 mm diameter aluminum spheres against 15 kg/m² Dyneema® HB26 at normal incidence (left) and impact of 20 mm FSP against 10 mm thick HB26 at normal incidence (right).

B. Method

The FSP material was modelled as Steel S-7 from the AUTODYN library using a linear EoS and the Johnson-Cook strength model [25]. The aluminum sphere was modelled using AL1100-O from the AUTODYN library that uses a shock EoS and the Steinburg Guinan strength model [26]. The master-slave contact algorithm was used to detect contact between the target and projectile.

The sub-laminate model with shock EoS was applied to the aluminum sphere hypervelocity impact series and 20 mm FSP ballistic impact series presented in Figure 8, the results of which are shown in Figure 9.

The sub-laminate model is shown to provide a significant improvement in predicting the experimental V₅₀ of 394 m/s for the FSP ballistic impacts (377 m/s) compared to the monolithic model (236 m/s).

The ballistic limit and residual velocity predicted with the sub-laminate model for the hypervelocity impact case are shown to be comparable with the original monolithic model.

For conditions closer to the ballistic limit, the sub-laminate model is shown to predict increased target resistance (i.e., lower residual velocity). For higher overmatch conditions, there is some small variance between the two approaches.

Now, regarding common handgun projectiles, the results look sobering. As the most widespread weapon in the world, the Kalashnikov rifle (AK-47) is a good example to compare ballistic trials and simulation results. A 7.62×39 mm full metal jacket (FMJ) projectile with a velocity of 700 m/s is used and the model is shown in Figure 3.

In Figure 10, a qualitative assessment of the bulge formation is made for the 11 mm panel. Prediction of bulge development is important as it is characteristic of the material wave speed and is also a key measure in defence applications, particularly in personal armour systems (i.e., vests and helmets). In the ballistic experiments, the 11 mm target panel resists the 7.62×39 mm FMJ projectile. But in both models, material fails and the projectile penetrates the plate.

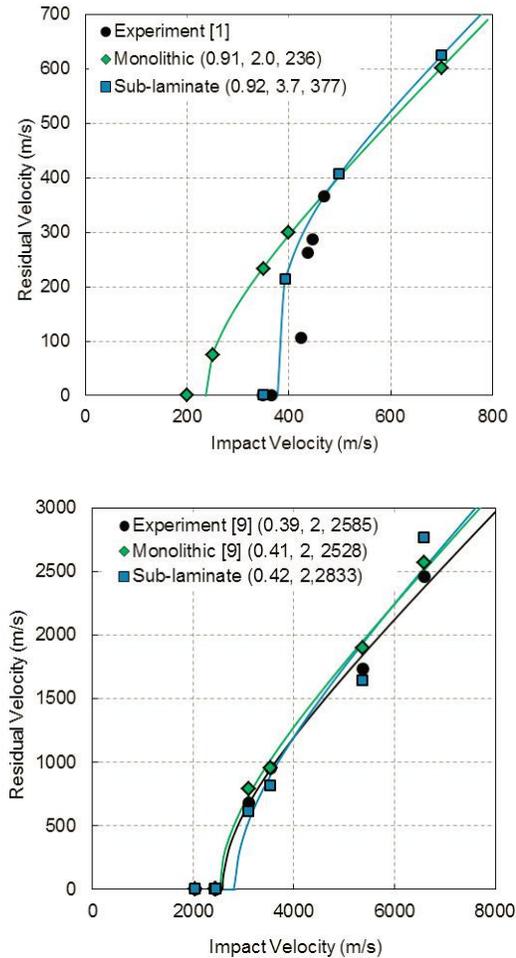


Figure 9. Comparison of the experimental results with the two numerical models for impact of 20 mm FSP against 10 mm thick Dyneema® HB26 at normal incidence (left), and impact of 6 mm diameter aluminium spheres against 15 kg/m² HB26 at normal incidence (right).

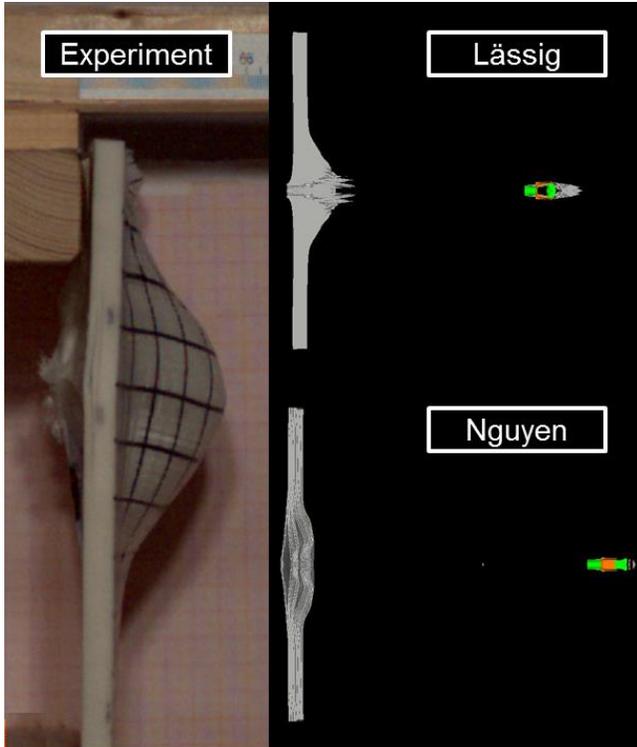


Figure 10. Bulge of a 11 mm target impact by a 7.62×39 mm FMJ projectile at 689 m/s (experiment) and the simulation results using both “state-of-the-art” models of Lässig and Nguyen 700 m/s.

There is no accurate reproduction of the bulge. The problem is a neglect of micro-structures. Fraction and fragmentation between the laminate layers cannot be described by homogeneous continuum models. These disadvantages are addressed in a very new and more representative model.

VI. NUMERICAL SIMULATION

As mentioned before, the ballistic tests are followed by computational modelling of the experimental set-up. Then, the experiment is reproduced using numerical simulations. Figure 3 shows a cross-section of the projectile and a CAD model. The geometry and observed response of the laminate to ballistic impact is approximately symmetric to the axis through the bullet impact point.

Numerical simulation of modern armor structures requires the selection of appropriate material models for the constituent materials and the derivation of suitable material model input data. The laminate system studied here is an ultra-high molecular weight polyethylene composite. Lead and copper are also required for the projectiles.

The projectile was divided into different parts - the jacket and the base - which have different properties and even different meshes. These elements have quadratic shape functions and nodes between the element edges. In this way, the computational accuracy, as well as the quality of curved model shapes increases. Using the same mesh density, the application of parabolic elements leads to a higher accuracy compared to linear elements (1st order elements).

Modelling of fiber-reinforced composites under impact is challenging because of the complexity of the material composition and the many failure modes it exhibits at different scales (fibrillation, intra- and inter-laminar failure, etc.) and impact regimes. For this reason, numerical simulation of impact using hydrocodes was exclusively performed for isotropic materials up until the late 1990’s. Since then, there have been many advances in modelling composites brought about by the introduction of more accurate constitutive models and modelling techniques. In general, fiber-reinforced composites can be modelled at three different scales, as shown in Figure 11:

- Micro-scale, where the individual fiber, matrix and (in some cases) the fiber-matrix interface is explicitly modelled;
- Meso-scale, where the properties of the individual plies that are homogenised in the principal directions are modelled and stacked together to produce a laminate; and
- Macro-scale, where the laminate is modelled as a continuum and the properties of the laminate are homogenised in the principal directions.

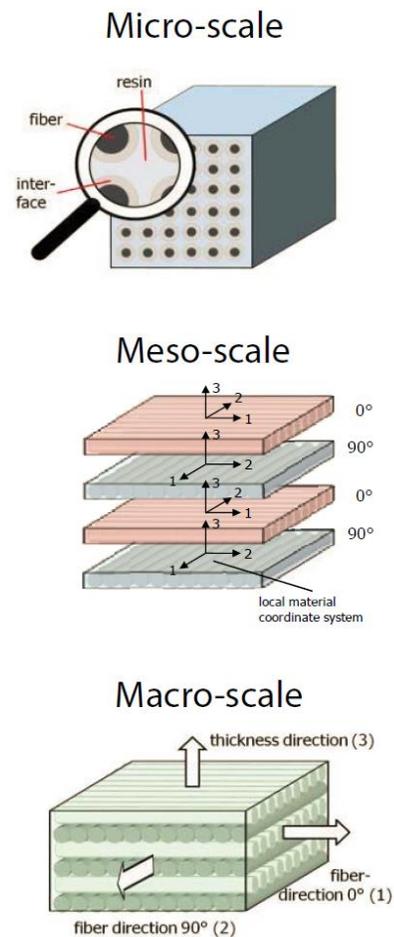


Figure 11. Micro, meso and macro mechanical model of fiber reinforced composites [27].

Modelling of fiber-reinforced composites at the micro-scale has several important advantages. This includes increased model fidelity, relatively simpler constitutive equations to describe the fiber, matrix and the interface, and characterisation tests that are relatively easy to perform. However, models at this scale require explicit modelling of every single fiber, matrix and the contact interface, which is extremely computationally expensive and not practical currently for typical engineering problems (see Figure 12).

While the meso-scale approach is far more computationally tractable compared to the micro-scale, models at this scale are still not practical for thick targets.

A. Modelling

Because of the discrepancies discussed in Section V, a new model was developed – a concept for the numerical simulation of fiber-reinforced plastics under impact loading. Here, the homogeneous continuum model is replaced. Alternating layers of fibers and matrix are used for the geometry. The layers are bonded and have different material models (see Figure 13).

The fiber layers apply anisotropic elasticity, no plasticity and anisotropic material failure (stress-dependent). The matrix layers use isotropic elasticity, von Mises plasticity and isotropic material failure (stress-dependent). To simulate the effect of delamination, principal stress failure is applied.

3D numerical simulations were performed of the full target and projectile, where both were meshed using 8-node hexahedral elements. The projectile was meshed with 9 elements across the diameter. The target is composed of sub-laminates that are one element thick, separated by a small gap to satisfy the master-slave contact algorithm (external gap in AUTODYN) and bonded together as previously discussed. The mesh size of the target is approximately equal to the projectile at the impact site. The mesh was then graded towards the edge, increasing in coarseness to reduce the computational load of the model. Since UHMW-PE composite has a very low coefficient of friction, force fit clamping provides little restraint. High speed video of ballistic impact tests typical showed clamp slippage upon impact. As such no boundary conditions were imposed on the target.

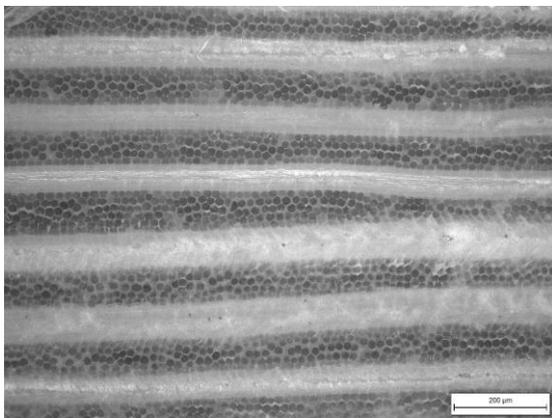


Figure 12. Cross section of a Dyneema® HB26 panel.

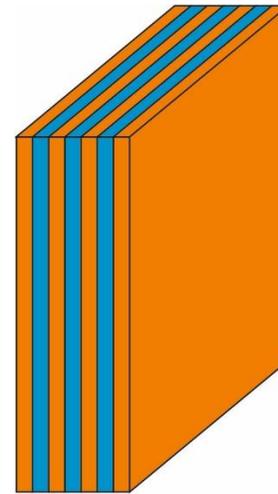


Figure 13. Geometry of a target plate: alternating layers of fibers and matrix are used in the computer model.

B. Simulation Results

The model developed in [16] was adjusted and the concept has been extended to different calibers and projectile velocities. Composite armor plates between 5.5 and 16.2 mm were tested in several ballistic trials and high-speed videos were used to analyze the characteristics of the projectile – before and after the impact. The simulation results with the modified model and a 7.62×39 mm bullet are shown in Figures 14, 15, and 16.

The deformation of the projectile is in good agreement with the experimental observation. Both delamination and fragmentation can be seen in the numerical simulation. Compared to the homogeneous continuum model, fractures can be detected easily. Subsequently, the results of experiment and simulation in the case of perforation were compared with reference to the projectile residual velocity. Here, only minor differences were observed. The results are summarized in Table I.

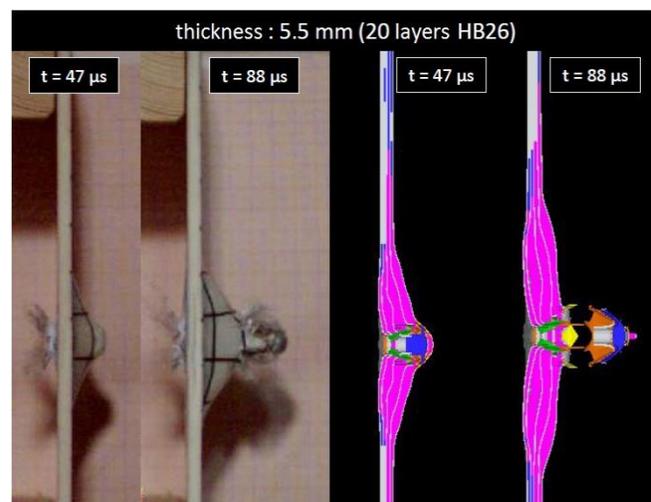


Figure 14. Effect of a 5.5 mm target impact by a 7.62×39 mm bullet at 686 m/s, 47 μs and 88 μs after the initial impact.

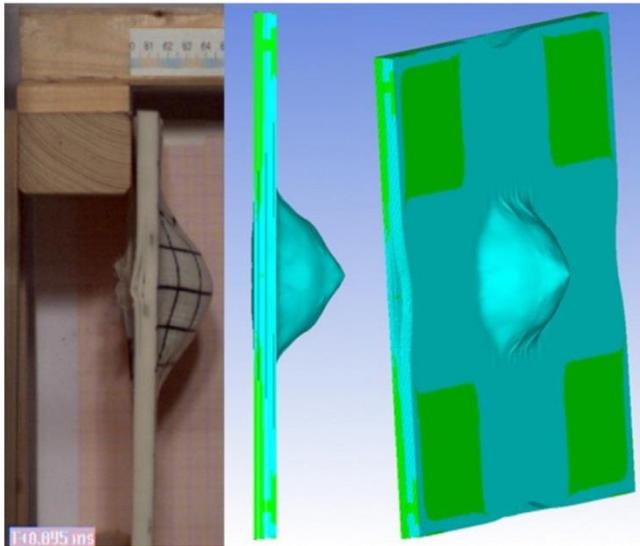


Figure 15. Effect of a 11.0 mm target impact by a 7.62×39 mm bullet at 682 m/s.

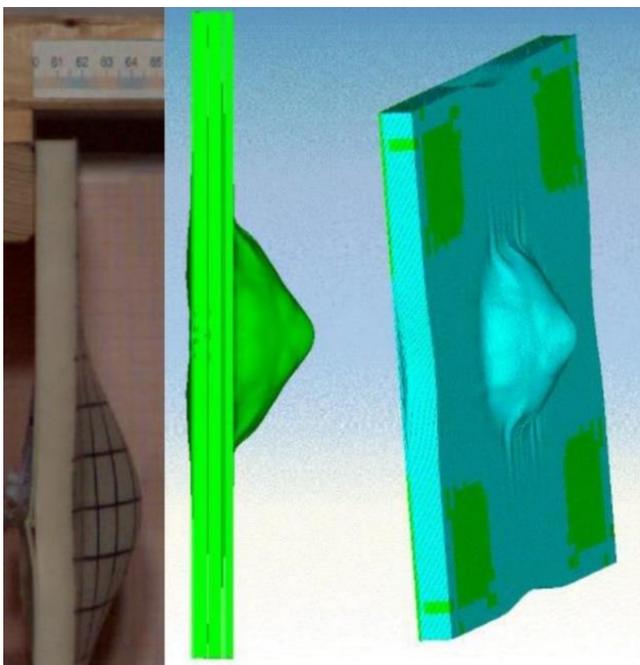


Figure 16. Effect of a 16.2 mm target impact by a 7.62×39 mm bullet at 679 m/s.

TABLE I. SUMMARY OF RESULTS FOR A 7.62×39 MM BULLET

Target Thickness	5.5 mm Residual Velocity	11.0 mm Residual Thickness	16.2 mm Residual Thickness
Experiment	604 m/s	6 mm	11 mm
Simulation	587 m/s	4 mm	10 mm

Even the simulation results with other calibers provide very good results. A common projectile is the 9×19 mm Parabellum. Under STANAG 4090, it is a standard cartridge for NATO forces as well as many non-NATO countries. According to the 2014 edition of Cartridges of the World, the 9×19 mm Parabellum is “the world’s most popular and widely used military handgun and submachinegun cartridge.” In addition to being used by over 60% of police in the U.S., the 9×19 mm pistols are more popular than revolvers. The simulation results with the modified model and a 9×19 mm bullet are shown in Figures 17, 18, and 19.

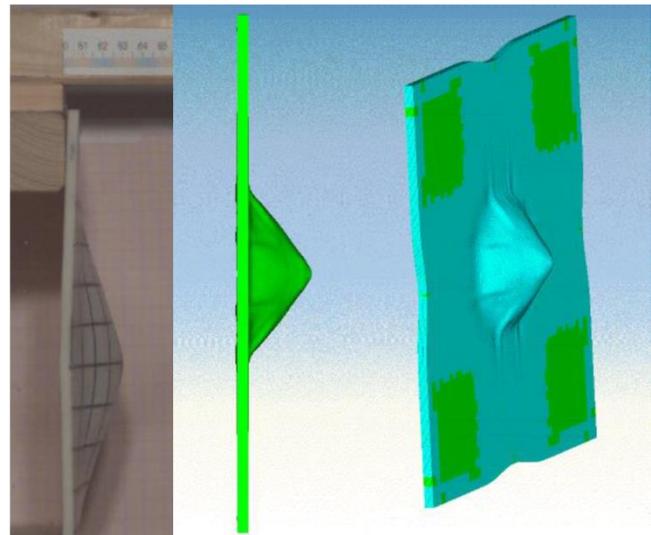


Figure 17. Effect of a 5.5 mm target impact by a 9×19 mm Parabellum bullet at 348 m/s.

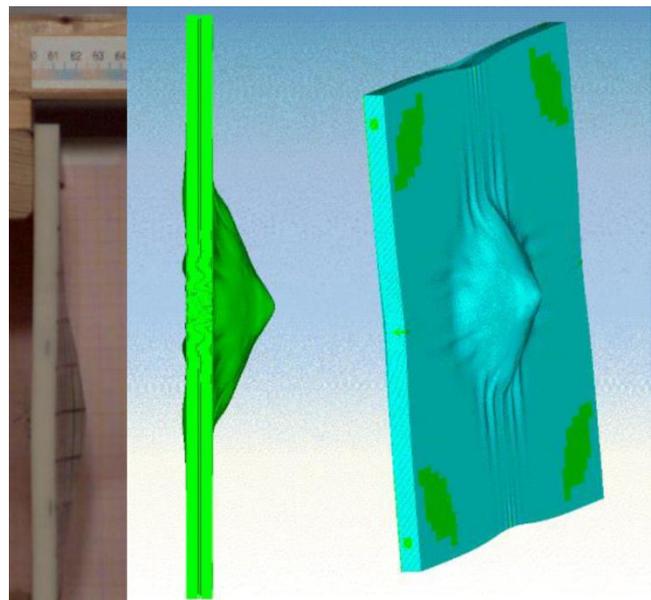


Figure 18. Effect of a 11.0 mm target impact by a 9×19 mm Parabellum bullet at 343 m/s.

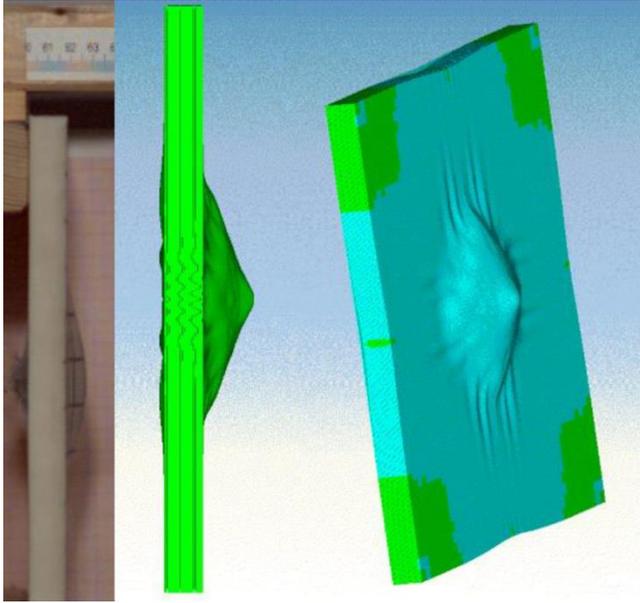


Figure 19. Effect of a 16.2 mm target impact by a 9×19 mm Parabellum bullet at 346 m/s.

The next step is to test the influence of meshing, based on Ramezani and Rothe [28]. The accuracy that can be obtained from any FEA model is directly related to the finite element mesh that is used. The finite element mesh is used to subdivide the CAD model into smaller domains called elements, over which a set of equations are solved. These equations approximately represent the governing equation of interest via a set of polynomial functions defined over each element. As these elements are made smaller and smaller, as the mesh is refined, the computed solution will approach the true solution. This process of mesh refinement is a key step in validating any finite element model and gaining confidence in the software, the model, and the results.

Early in the analysis process, it makes sense to start with a mesh that is as coarse as possible – a mesh with very large elements. A coarse mesh will require less computational resources to solve and, while it may give a very inaccurate solution, it can still be used as a rough verification and as a check on the applied loads and constraints.

After computing the solution on the coarse mesh, the process of mesh refinement begins. In its simplest form, mesh refinement is the process of resolving the model with successively finer and finer meshes, comparing the results between these different meshes. This comparison can be done by analyzing the fields at one or more points in the model or by evaluating the integral of a field over some domains or boundaries. Table II summarizes the four different mesh sizes. A Hewlett-Packard (HP) ProLiant DL380p G8 Server is used for all calculations. By comparing these scalar quantities, it is possible to judge the convergence of the solution with respect to mesh refinement. The results are shown in Figure 20. It should be noted that an explicit modelling of the individual fibers is not an option, since the computational effort would go beyond the scope of modern server systems.

TABLE II. COMPARISON OF THE CALCULATION TIME

Mesh Size	Number of Elements over Edge Length	Calculation Time [h]
Coarse	25	0.5
Medium	50	4
Fine	75	25
Very Fine	100	268

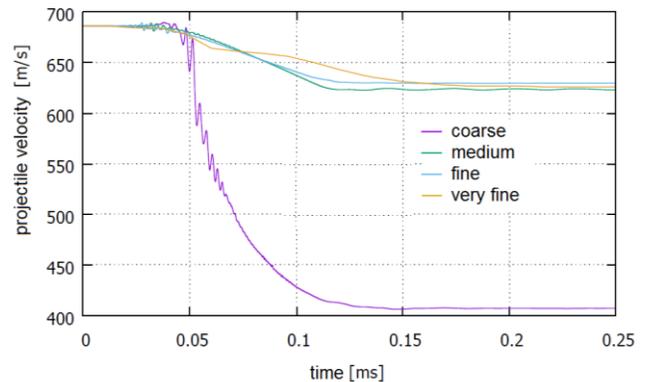


Figure 20. Convergence analysis: comparison of the projectile velocity (7.62×39 mm bullet) at different mesh sizes.

VII. CONCLUSIONS

The material model developed in [16] has some shortcomings regarding the simulation of handgun projectiles (e.g., 7.62×39 mm). Although previously found to provide accurate results for hypervelocity impact of aluminum spheres, the existing model and dataset was found to significantly underestimate the composite performance under impact conditions driven by through-thickness shear performance (ballistic impact of fragment simulating projectiles). The model was found to exhibit premature through thickness shear failure as a result of directional coupling in the modified Hashin-Tsai failure criterion and the large discrepancy between through-thickness tensile and shear strength of UHME-PE composite. As a result, premature damage and failure was initiated in the through-thickness shear direction leading to decreased ballistic performance. By de-coupling through-thickness tensile failure from the failure criteria and discretizing the laminate into a nominal number of kinematically joined sub-laminates through the thickness, progresses in modelling the ballistic response of the panels was improved.

A major difficulty in the numerical simulation of fiber composites under impact is the detection of failure processes between fiber and matrix elements as well as between the individual laminate layers (delamination). One promising approach is the use of "artificial" inhomogeneities on the macroscale.

This paper is based on Ramezani and Rothe [29]. New approaches make it possible to increase the accuracy of the

simulation results. The previous concept was not valid for all projectiles/calibers. An alternative model has been developed to overcome these difficulties. Using sub-laminates and inhomogeneities on the macroscale, the model does not match the real microstructure, but allow a more realistic description of the failure processes mentioned above. The numerical model proposed a sub-laminate discretisation of the laminate in order to better model delamination failure. For this to occur, the sub-laminates are joined together using bonded contacts where failure is initiated based on a combined normal and shear stress criterion. Such approach is known to be mesh-dependent and has been replaced by fracture energy based failure (mode I, II, and III) [30]. Most finite element models of fibre-reinforced composite today capture interlaminar failure through zero thickness cohesive elements which accounts for both damage and fracture mechanics (using traction-separation laws), though this is not available in the hydrocode used in this work.

This work also demonstrated how a small number of well-defined experiments can be used to develop, calibrate, and validate solver technologies used for simulating the impact of projectiles on complex armor systems and composite laminate structures. Ballistic trials can be used as the basis of an iterative optimization process. Numerical simulations are a valuable adjunct to the study of the behavior of metals subjected to high-velocity impact or intense impulsive loading. The combined use of computations, experiments and high-strain-rate material characterization has, in many cases, supplemented the data achievable by experiments alone at considerable savings in both cost and engineering man-hours.

VIII. FUTURE WORK

Generally, the field of ballistic and penetration mechanics is extensive due to the unlimited combination of targets and threats. The response of targets is different depending on the projectile size, geometry, material and impact velocity. The scope of this work was restricted to the most common threats (where FSPs are used as a representative surrogate), however, understanding the penetration and failure mechanisms of the material impacted by different projectiles (spherical and ogive) can also be valuable. This work also only considered normal impacts (the worst-case scenario), but attacks experienced on the front line are almost always at an oblique angle (large or small). Understanding how UHMW-PE composite responds to obliquity is important and deserves attention.

REFERENCES

- [1] A. Ramezani and H. Rothe, "A New Approach to Modelling Fiber-Reinforced Plastics for Hydrocode Analysis - Experimental Model Validation of Composite Materials Under Ballistic Impact," The Ninth International Conference on Advances in System Simulation (SIMUL 2017) IARIA, Oct. 2017, pp. 41-50, ISBN 978-1-61208-594-4.
- [2] R&G Faserverbundwerkstoffe GmbH, Handbuch Faserverbundwerkstoffe. D-71111 Waldenbuch, 2009.
- [3] J. Van Dingenen, "Gel-spun high-performance polyethylene fibres," in *High Performance Fibres*, pp. 62–92, 2001.
- [4] B. Russell, K. Karthikeyan, V. Deshpande, and N. Fleck, "The high strain rate response of ultra-high molecular-weight polyethylene: From fibre to laminate," in *International Journal of Impact Engineering* (60), pp. 1–9, 2013.
- [5] L. H. Nguyen, *The Ballistic Performance of Thick Ultra High Molecular Weight Polyethylene Composite*. RMIT University, 2015.
- [6] P. M. Cunniff, "Dimensionless parameters for optimization of textile based body armor systems," in *18th International Symposium on Ballistics*, pp. 1303–1310, 1999.
- [7] J. Zukas, *Introduction to hydrocodes*. Elsevier Science, 2004.
- [8] G.-S. Collins. *An Introduction to Hydrocode Modeling*. Applied Modelling and Computation Group, Imperial College London, 2002.
- [9] E. Madenci and I. Guven, "The Finite Element Method and Applications in Engineering Using ANSYS," Springer Verlag, February 2015.
- [10] H.-B. Woyand, *FEM with CATIA V5*. J. Schlembach Fachverlag, 2007.
- [11] S. J. Hiermaier, *Structures Under Crash and Impact*. Springer Verlag, 2008.
- [12] R. F. Stellingwerf and C. A. Wingate, "Impact Modeling with Smooth Particle Hydrodynamics," *International Journal of Impact Engineering*, vol. 14, pp. 707–718, Sep. 1993.
- [13] X. Quan, N. K. Birnbaum, M. S. Cowler, and B. I. Gerber, "Numerical Simulations of Structural Deformation under Shock and Impact Loads using a Coupled Multi-Solver Approach," 5th Asia-Pacific Conference on Shock and Impact Loads on Structures, Hunan, China, Nov. 2003, pp. 152-161.
- [14] ANSYS Inc. Available Solution Methods. [Online]. Available from: <http://www.ansys.com/Products/Simulation+Technology/Structural+Analysis/Explicit+Dynamics/Features/Available+Solutions+Methods> [retrieved: April, 2014]
- [15] P. Fröhlich, "FEM Application Basics," Vieweg Verlag, September 2005.
- [16] D. B. Segala and P. V. Cavallaro, "Numerical investigation of energy absorption mechanisms in unidirectional composites subjected to dynamic loading events," in *Computational Materials Science* 81, pp. 303–312, 2014.
- [17] S. Chocron et al., "Modeling unidirectional composites by bundling fibers into strips with experimental determination of shear and compression properties at high pressures," in *Composites Science and Technology* 101, pp. 32–40, 2014.
- [18] C. J. Hayhurst, S. J. Hiermaier, R. A. Clegg, W. Riedel, and M. Lambert, "Development of material models for nextel and kevlar-epoxy for high pressures and strain rates," in *International Journal of Impact Engineering* 23, pp. 365–376, 1999.
- [19] R. A. Clegg, D. M. White, W. Riedel, and W. Harwick, "Hypervelocity impact damage prediction in composites: Part I—material model and characterisation," in *International Journal of Impact Engineering* 33, pp. 190–200, 2006.
- [20] W. Riedel, H. Nahme, D. M. White, and R. A. Clegg, "Hypervelocity impact damage prediction in composites: Part II—experimental investigations and simulations," in *International Journal of Impact Engineering* 33, pp. 670–80, 2006.
- [21] M. Wicklein, S. Ryan, D. M. White, and R. A. Clegg, "Hypervelocity impact on CFRP: Testing, material modelling, and numerical simulation," in *International Journal of Impact Engineering* 35, pp. 1861–1869, 2008.

- [22] ANSYS. AUTODYN Composite Modelling Release 15.0. [Online]. Available from: <http://ansys.com/> [retrieved: May, 2017].
- [23] T. Lässig et al., "A non-linear orthotropic hydrocode model for ultra-high molecular weight polyethylene in impact simulations," in *International Journal of Impact Engineering* 75, pp. 110–122, 2015.
- [24] C. J. Hayhurst, J. Leahy, J., M. van der Jagt-Deutekom, and P. Kelly, "Development of Material Models for Numerical Simulation of Ballistic Impact onto Polyethylene Fibrous Armour," in *Personal Armour Systems Symposium*, pp. 5–8, 2000.
- [25] K. Herlaar, M. van der Jagt-Deutekom, and M. Jacobs, "Finite element simulations and experiments of ballistic impacts on high performance PE composite material," in *22nd International Symposium on Ballistics*, pp. 1040–1047, 2005.
- [26] C. Ong, C. Boey, R. S. Hixson, and J. O. Sinibaldi, "Advanced layered personnel armor," in *International Journal of Impact Engineering* 38(5), pp. 369–383, 2011.
- [27] U. Heisserer and H. Van der Werff, "The relation between Dyneema fiber properties and ballistic protection performance of its fiber composites," in *15th International Conference on Deformation, Yield and Fracture of Polymers* 3, pp. 242–246, 2012.
- [28] A. Ramezani and H. Rothe, "Influences of Meshing and High-Performance Computing towards Advancing the Numerical Analysis of High-Velocity Impacts," in *International Journal on Advances in Systems and Measurements*, Volume 9, Numbers 3 & 4, 2016, ISSN 1942-261x , S.142-153.
- [29] A. Ramezani and H. Rothe, "Numerical Simulation and Experimental Model-Validation for Fiber-Reinforced Plastics Under Impact Loading - Using the Example of Ultra-High Molecular Weight Polyethylene," *The Eighth International Conference on Advances in System Simulation (SIMUL 2016) IARIA*, Aug. 2016, pp. 17-25, ISBN 978-1-61208-442-8.
- [30] Y. Mi, M. Crisfield, G. Davies, and H. Hellweg., "Progressive delamination using interface elements," in *Journal of Composite Materials*, 32(14), pp. 1246–1272, 1998.

A Framework for Optimizing Simulation Model Validation & Verification

Bill Roungas
Alexander Verbraeck

Department of Multi Actor Systems
Delft University of Technology
Delft, The Netherlands
Email: v.roungas@tudelft.nl,
a.verbraeck@tudelft.nl

Sebastiaan Meijer

Department of Health Systems Engineering
KTH Royal Institute of Technology
Huddinge, Sweden
Email: sebastiaan.meijer@sth.kth.se

Abstract—Thirty years of research on validation and verification have returned a plethora of methods and statistical techniques through methodological and case studies. It is, however, this abundance of methods and techniques that poses a major challenge. Due to time and budget constraints, it is impossible to apply all the available methods and techniques in a single study, and as such a careful selection has to be made. This paper builds on two assumptions: a) simulations, real-world systems, methods, and techniques can be defined on the basis of different characteristics and b) certain methods and techniques are more suitable than others for different kinds of simulation studies. The present study aims at identifying the specific characteristics that make certain methods and techniques more effective and more efficient than others, when juxtaposing these with the simulations' and systems' different characteristics. The conclusion will advance a methodology for choosing the most appropriate methods and techniques for validating and/or verifying a simulation.

Keywords—simulation; validation; verification; method selection.

I. INTRODUCTION

Back in 1972, based on Forrester's work [1][2], Meadows et al. [3][4] introduced World 3, a simulation of the world for the years 1900-2100. The purpose of the simulation model was to project the dynamic behavior of population, capital, food, non-renewable resources, and pollution. The model's forecast was that during the contemplated two centuries the world will experience a major industrial collapse, which will be followed by a significant decrease in human population. The model became very popular especially because of the increasing interest in environmental degradation encountered because of human activities [5]. Even though the model gained support for being "of some use to decision makers" [4] and generated the spark for many later global models, it had several shortcomings, for which it received a lot of criticism [6]. In turn, this criticism raised the question of whether, and to what extent, such simulation models are validated and verified. This is just one example of the notion that validation and verification (V&V) is a fundamental part of a simulation study [7].

The term V&V is used to characterize two relatively different approaches that almost always go hand by hand, namely validation and verification. Validation is this phase of a study that ensures that the simulation imitates the underline system, to a greater or lesser extent, and in any case satisfactorily [8], or in layman terms validation addresses the question of whether the built model is the also the "right" one [9]. On the other hand, verification is the phase of the study that ensures

that the model and its implementation are correct [10], or in layman terms verification addresses the question of whether the model was built in the "right" way [9]. V&V has become a well-researched field with a significant amount of produced literature and commercial case studies. The large number of V&V methods and statistical techniques created or adopted by this wide range of research, is the greatest impediment to designing a V&V study.

The predetermined budget of a simulation study usually limits the amount of time and resources that can be spent on V&V. Additionally, the nature and the diverse characteristics of simulations limit the number of V&V methods and statistical techniques that are applicable to each simulation. In other words, not all V&V methods (hereinafter referred to as methods) and V&V-applicable statistical techniques (hereinafter referred to as techniques) are suitable for every simulation. To the best of our knowledge, a taxonomy for characterizing methods and techniques and, subsequently, matching them with different simulations does not exist. Therefore, the research question that this study will address is:

How can the selection of V&V methods and V&V-applicable statistical techniques given the simulation and the real-world system at hand be optimized as to be more time efficient and rigorous?

This paper aims at identifying the majority of the available methods and techniques in order to classify them on the basis of their different characteristics and on whether they can be used to validate or verify a simulation, and eventually match them with characteristics of simulation models.

In Section II, a literature analysis on methods and techniques, simulation properties, and simulation study phases is conducted. In Section III, a methodology towards developing a framework for simulation V&V method and statistical techniques selection is proposed. In Section IV, a case study is presented to illustrate how the proposed framework can actually be implemented. Finally, in Section V, the future potential extensions of the framework are presented and final remarks are made.

II. LITERATURE ANALYSIS

In this section, a 3-step literature analysis is presented. The initial hypothesis of this study is that simulations exhibit certain properties that influence the effectiveness and applicability of methods and techniques. Therefore, the 3 steps of the literature analysis are the following:

Step 1: Identification of methods and techniques.

Step 2: Identification of simulations' properties potentially influencing the selection of methods and identification of simulations' and systems' characteristics potentially influencing the selection of techniques.

Step 3: Identification of the phases of a simulation study.

A. Step 1: V&V methods and statistical techniques

Methods are different in many aspects; some methods are strictly mathematical whereas others accommodate the more qualitative aspects of simulations, etc. Balci [11] identified more than 70 methods, which in turn categorized into four categories: informal, static, dynamic, and formal. Balci's [11] list is the most accurate representation of the body of work on methods and, even to date, is considered as the most extensive one. This paper adopts the list in reference - but not the categorization - and goes as far as to propose a new classification of methods.

On the other hand, numerous techniques have been proposed throughout the years, a subset of which are applicable in V&V studies. Moreover, techniques can be characterized in various ways, e.g., depending on the input they require (numerical, categorical etc.), or the purpose they are used for (goodness-of-fit, time series etc.).

In Section II-A1 and Section II-A2, the identified methods and techniques are listed, respectively, along with a brief definition for each one of them.

1) V&V Methods:

Acceptance Testing: Acceptance Testing is testing the model using the actual hardware and data to determine whether all the specified requirements are satisfied [12].

Alpha Testing: Alpha Testing is the operational testing of the alpha version of the model in a department within the company, yet not the one involved with the model development [13].

Assertion Checking: Assertion Checking checks what is happening as opposed to what the modeler assumes is happening thus detecting potential errors in the model [11].

Audit: An Audit is undertaken to assess how adequately the simulation study is conducted with respect to established plans, policies, procedures, standards and guidelines. The audit also seeks to establish traceability within the simulation study [11].

Beta Testing: Beta Testing is the operational testing of the beta version of the model under realistic field conditions [14].

Bottom-Up Testing: Bottom-Up Testing is testing each sub-model, when the model is developed with a bottom-up development strategy, and once every submodel belonging to the same parent is finished and tested, then these submodels are integrated and tested again [11].

Cause-Effect Graphing: Cause-Effect Graphing aids in selecting, in a systematic way, a high-yield set of test cases and it is effective in pointing out incompleteness and ambiguities in the specification [15].

Comparison Testing: Comparison Testing is testing the different versions of the same simulation model [16].

Compliance Testing: Compliance Testing tests how accurately different levels of access authorization are provided, how closely and accurately dictated performance requirements are satisfied, how well the security requirements are met, and how properly the standards are followed [17]. It consists of the following techniques:

- 1) Authorization Testing, which tests how accurately and properly different levels of access authorization

are implemented in the model and how properly they comply with the rules and regulations [12].

- 2) Performance Testing, which tests whether (a) all performance characteristics are measured and evaluated with sufficient accuracy, and (b) all established performance requirements are satisfied [12].
- 3) Security Testing, which tests whether all security procedures are correctly and properly implemented in conducting a simulation study [12].
- 4) Standards Testing, which substantiates that the simulation model is developed with respect to the required standards, procedures, and guidelines [11].

Control Analysis: Control Analysis analyzes the control characteristics of the model. It consists of the following techniques:

- 1) Calling Structure Analysis, which is used to assess model accuracy by identifying who calls whom and who is called by whom [14].
- 2) Concurrent Process Analysis, in which model accuracy is assessed by analyzing the overlap or concurrency of model components executed in parallel or as distributed [18].
- 3) Control Flow Analysis, which requires the development of a graph of the model where conditional branches and model junctions are represented by nodes and the model segments between such nodes are represented by links [13].
- 4) State Transition Analysis, which requires the identification of a finite number of states the model execution goes through and shows how the model transitions from one state to another [19].

Data Analysis: Data Analysis ensures that (1) proper operations are applied to data objects (e.g., data structures, event lists, linked lists), (2) the data used by the model are properly defined, and (3) the defined data are properly used [12]. It consists of the following techniques:

- 1) Data Dependency Analysis, which involves the determination of what variables depend on what other variables [20].
- 2) Data Flow Analysis, which is used to assess model accuracy with respect to the use of model variables [21].

Debugging: Debugging identifies errors causing the model to fail and changes the model accordingly in order to correct these errors [20].

Desk Checking: Desk Checking is when a person other than the modeler thoroughly examines the model to ensure correctness, completeness, consistency and unambiguity [13].

Documentation Checking: Documentation Checking ensures accuracy and up-to-date description of the model logic and its results [11].

Execution Testing: Execution Testing collects and analyzes execution behavior data in order to reveal model representation errors. It consists of the following techniques:

- 1) Execution Monitoring, which examines low-level information about activities and events taking place during model execution [11].
- 2) Execution Profiling, which examines high-level information about activities and events taking place during model execution [11].
- 3) Execution Tracing, which tracks line-by-line the execution of a model [11].

Face Validation: In Face Validation, people knowledgeable about the system under study subjectively compare model and system behaviors and judge whether the model and its results are reasonable [22].

Fault/Failure Analysis: Fault/Failure Analysis determines if any faults or failures can logically occur and in what context and under what conditions [14].

Fault/Failure Insertion Testing: Fault/Failure Insertion testing inserts an fault or failure into the model and observes whether the model will behave in the expected invalid manner [11].

Field Testing: Field Testing executes the model in an operational situation for the purpose of collecting information regarding the model validation [23].

Functional (Black-Box) Testing: Functional Testing ignores the internal mechanism(s) of the model and focuses on the generated outputs based on specific input and execution conditions [24].

Graphical Comparisons: In Graphical Comparison, graphs produced from the model are compared to graphs produced by the real-world system under study, in order to detect similarities and differences between the two [14].

Induction: Induction asserts that if every step a model follows is valid and the model terminates, then the model is valid [11]. Induction as a term can be found in many fields, like mathematics in which case it is a tool for directly proving theorems. In simulation model validation, where absolute validity does not exist [25], induction should more correctly be referred to as inductive reasoning, which is based on one or more inductive arguments, and the conclusions are not considered as the absolute truth but rather a strong evidence [26].

Inference: Inference is similar to Induction; it is a mental process by which one proposition is arrived at and affirmed on the basis on one or more other propositions assumed as the starting point of the process [26].

Inspections: Inspection is a five phase procedure conducted by four to six people. The phases include not only a validation phase but also suggestions for improvements and a follow-up [27].

Interface Analysis: Interface Analysis consists of the following techniques:

- 1) Model Interface Analysis, which is conducted to examine the (sub)model-to-(sub)model interface and determine if the interface structure and behavior are sufficiently accurate [11].
- 2) User Interface Analysis, which is conducted to examine the user-model interface and determine if it is human engineered so as to prevent occurrences of errors during the user's interactions with the model [11].

Interface Testing: Interface Testing consists of the following techniques:

- 1) Data Interface Testing, which assesses the accuracy of data inputted into the model or outputted from the model during its execution [14].
- 2) Model Interface Testing, which detects model representation errors caused due to interface errors [11].
- 3) User Interface Testing, which deals with the assessment of the interactions between the user and the model, and detects errors associated with those [27].

Lambda (λ) Calculus: λ -calculus is a mathematical tool for formally defining systems [28]. λ -calculus offers function that can be translated into validation rules.

Logical Deduction: Logical Deduction, also known as Deductive Reasoning, is similar to Induction but the conclusions are considered as logically true, or valid, if every step of the model is valid and the model terminates [26].

Object-Flow Testing: Object-Flow Testing assesses model accuracy by exploring the life cycle of an object during the model execution [11].

Partition Testing: Partition Testing, also known as equivalent class partitioning, partitions the model into functional representatives (partitions), assuming that all elements within each partition bear the same properties, and then, by selecting a representative element from each partition, each partition and subsequently the model is validated, thus eliminating the need for exhaustive validation [29].

Predicate Calculus: Predicate Calculus quantifies simple relationships (predicates) using boolean variables. Since, the model can be defined based on predicates, then its validation can be performed by manipulating these predicates [11]. Similarly to Deduction, Predicate Calculus' conclusions are logically true or valid.

Predicate Transformations: Predicate Transformations, or more formally known as Predicate Transformer Semantics, show that systems (in this case a simulation model) can achieve their goals, i.e., they are valid. Predicate Transformations associate a pre-condition to any post-condition, or in other words transform model output states into all model input states, thus providing the basis for proving model correctness [30].

Predictive Validation: In Predictive Validation, the model executes with past input data and the results are then compared with data from the real system [31].

Product Testing: Product Testing is a preparatory step for the Acceptance Testing, in which all requirements specification are tested in the same way as in the Acceptance Testing, with the only difference being that the test takes place within the development team whereas Acceptance Testing takes place at the client's premises [27].

Proof of Correctness: A Proof of Correctness expresses the model in a precise notation and then proves that the model terminates and thus satisfies the requirements specification with sufficient accuracy [32].

Regression Testing: Regression Testing ensures that correcting errors in the model during the validation process do not create new errors or adverse side-effects [11].

Reviews: Reviews are similar to an inspection but the review team also involves managers. Reviews are intended to give management and study sponsors evidence that the model development is being conducted according to the study objectives [12].

Semantic Analysis: Semantic Analysis attempts to determine the modeler's intent in writing the code [33].

Sensitivity Analysis: In Sensitivity Analysis, selected variables in the model are given different values (within a predetermined range) in order to observe the behavior of the model with regards to these changes [23].

Special Input Testing: Special Input Testing assesses model accuracy by subjecting the model in a variety of inputs and consists of the following techniques:

- 1) Boundary Value Testing, which tests the boundary values of the input and output equivalence classes (a set of values that bear similar characteristics and one value can act as a representative for the whole set)

- [15].
- 2) Equivalence Partitioning Testing, which tests the model by partitioning input data into equivalence classes [12].
 - 3) Extreme Input Testing, which tests the model based on extreme input values (minimum, maximum, or a mixture of those) [11].
 - 4) Invalid Input Testing, which tests the model using incorrect input data [11].
 - 5) Real-Time Input Testing, which tests the model using real-time input data from the real system [11].
 - 6) Self-Driven Input Testing, which test the model by executing it under input data randomly sampled from probabilistic models representing random phenomena of a real system [11].
 - 7) Stress Testing, which tests the model by subjecting it into heavy loads, like large volumes of data, intense activity over a short time span etc [15].
 - 8) Trace-Driven Input Testing, which tests the model by executing it under input trace data collected from a real system [11].

Structural (White-box) Testing: Structural Testing is used to evaluate the internal structure of the model and consists of the following techniques:

- 1) Branch Testing, which tests the model under test data in order to execute as many branch alternatives as possible [13].
- 2) Condition Testing, which tests the model under test data in order to execute as many logical conditions as possible [11].
- 3) Data Flow Testing, which tests the model by using the control flow graph as to explore sequences of events related to the status of data structures and to examine data-flow anomalies [13].
- 4) Loop Testing, which tests the model under test data in order to execute as many loop structures as possible [16].
- 5) Path Testing, which tests the model under test data in order to execute as many control flow paths as possible [13].
- 6) Statement Testing, which which tests the model under test data in order to execute as many statements as possible [13].

Structural Analysis: Structural Analysis is used to examine the model structure and to determine if it adheres to structured principles [11].

Submodel/Module Testing: Submodel/Module Testing is a top-down form of testing in which is submodel is tested against its corresponding subsystem [11].

Symbolic Debugging: Symbolic Debugging is a verification method in which the use of “breakpoints” allows for a direct manipulation of the model execution while viewing the model at the source code level [11].

Symbolic Evaluation: Symbolic Evaluation assesses model accuracy by executing the model using as an input symbolic values and not the actual data values [34].

Syntax Analysis: Syntax Analysis assures that the mechanics of the programming language are applied correctly [13].

Top-Down Testing: In Top-Down Testing, the model testing starts from the submodels at the highest level and moves downwards into the base submodels [35].

Traceability Assessment: Traceability Assessment matches, one to one, the elements of one form of the model to another [14].

Turing Test: In a Turing Test, experts are presented with two sets of output data, i.e., the model and reality, and without knowing which one is which, they are asked to differentiate the two [36].

Visualization/Animation: In Visualization/Animation, the model is tested by observing different graphs of the internal or external behavior of the model [11].

Walkthroughs: Walkthroughs are used to detect and document faults. Whilst They are similar to an Inspection but less time consuming, they have fewer phases [15].

2) *Statistical Techniques:* In the statistical formulas shown in this section, wherever M and R are used as subscripts, they denote that the particular variable refers to the model or reality respectively. Moreover, unless explicitly stated, n with the appropriate subscript denotes the respective sample size.

t-Test: The t-Test, also known as Student’s t-test, is a statistical hypothesis test, which determines whether the mean of a variable is significantly different from a constant value (one-sample test) or whether the mean of two variables is significantly different (two-sample test) [37]. The most common usage of t-test in simulation model V&V is the two-sample test (Model and Reality) with unequal sample sizes and variances. The latter is also known as Welch t-test [38] and its formula is:

$$t = \frac{\bar{X}_M - \bar{X}_R}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_R^2}{n_R}}} \quad (1)$$

where \bar{X} and s are the mean and variance respectively. The t-test is one of the most commonly used tests for the equality of means between model and reality.

Hotelling’s T^2 Test: Hotelling’s T^2 test is a generalization of the t-test for multivariate hypothesis testing [39]. As it is the case with t-test, Hotelling’s T^2 test can also be used for one- or two-sample testing. Its formula for the two-sample test is:

$$T^2 = (\bar{X}_M - \bar{X}_R)' \left\{ S_p \left(\frac{1}{n_M} + \frac{1}{n_R} \right) \right\}^{-1} (\bar{X}_M - \bar{X}_R) \quad (2)$$

where

$$X_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, i = \{M, R\} \quad (3)$$

$$S_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ij} - \bar{x}_i)(X_{ij} - \bar{x}_i)' \quad (4)$$

$$S_p = \frac{(n_M - 1)S_M + (n_R - 1)S_R}{n_M + n_R - 2} \quad (5)$$

Analysis of Variance (ANOVA): ANOVA is a collection of statistical techniques for testing mean equality between three or more datasets [40]. It is similar to multiple two-sample t-tests but less prone to a Type I error. The most popular ANOVA test is the F-Test. In a nutshell, the F-Test is the ratio of the variability between the datasets to the variability within each dataset [41]. The formula is:

$$F = \frac{\sum_{i=1}^K n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1)}{\sum_{i=1}^K \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 / (N - K)} \quad (6)$$

where Y_i is the average of the i^{th} dataset, \bar{Y} the overall average of the data, K the number of datasets, Y_{ij} the j^{th} observation of the i^{th} dataset, and N the total sample size.

Multivariate Analysis of Variance (MANOVA): MANOVA is similar to ANOVA but for cases where the dependent variables are more than one [42]. One of the most popular MANOVA tests is the Samuel Stanley Wilks' statistic, which is a summary based on the eigenvalues λ_p of the A matrix ($A = \sum_M * \sum_{res}^{-1}$), where \sum_M is the model variance matrix and \sum_{res} the error variance matrix. Wilks' formula is:

$$\Lambda_{Wilks} = \prod_{1..p} (1/(1 + \lambda_p)) = \det\left(\sum_{res}\right) / \det\left(\sum_{res} + \sum_M\right) \tag{7}$$

and is distributed as Λ .

Simultaneous Confidence Intervals: Balci and Sargent [43] proposed the validation method of simultaneous confidence intervals (sci) for simulation models with multiple outputs. The sci are formed by the confidence intervals of each model output. They described three approaches for calculating the sci and choosing one approach over the others depends on whether the model is self- or trace-driven. In other words, the choice of the approach depends on whether the model's input data are coming from the same population as the system's input data but they are different or whether the model's input data are exactly the same as the system's.

Factor Analysis: Using factor analysis, p observed random variables can be expressed as linear functions of m ($m < p$) random variables, also called common factors, along with an error [44]. If $x = \{x_1, x_2, \dots, x_p\}$ are the observed variables, $f = \{f_1, f_2, \dots, f_m\}$ the common factors, and $e = \{e_1, e_2, \dots, e_p\}$ the error, then there exists a

$$K = \begin{bmatrix} \kappa_{11} & \kappa_{12} & \dots & \kappa_{1m} \\ \kappa_{21} & \kappa_{22} & \dots & \kappa_{2m} \\ \dots & \dots & \dots & \dots \\ \kappa_{p1} & \kappa_{p2} & \dots & \kappa_{pm} \end{bmatrix} \tag{8}$$

so $x = Kf + e$.

Principal Component Analysis (PCA): The idea behind PCA is that if there is a large number (p) of random correlated variables, orthogonal transformation can be used to convert these variables into a significantly smaller number (m) of uncorrelated variables, called principal components [44]. PCA is similar to factor analysis, and is often considered to be a method of factor analysis. Despite their similarities, PCA and factor analysis are different in the sense that PCA concentrates on the diagonal elements of the covariance matrix, i.e., the variances, whereas the factor analysis focuses on the non-diagonal elements. In mathematical terms, PCA can be defined as follows:

$$f_1 = a'_1 x = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = \sum_{j=1}^p a_{1j}x_j \tag{9}$$

$$f_2 = a'_2 x = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p = \sum_{j=1}^p a_{2j}x_j \tag{10}$$

⋮

$$f_m = a'_m x = a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mp}x_p = \sum_{j=1}^p a_{mj}x_j \tag{11}$$

where f is the m principal components, a' is a transposed vector of constants, and x is the p independent variables. It should be noted that PCA is particularly useful when $m \ll p$. **Kolmogorov-Smirnov Test:** The Kolmogorov-Smirnov test (K-S test) is a non-parametric goodness-of-fit test that it can be one-sample, i.e., test whether a sample is distributed according to a known theoretical distribution (e.g., normal, binomial etc.), or two-sample, i.e., test whether two different samples are drawn from the same empirical distribution [45]. In simulation model V&V, the two-sample K-S test is the most common, i.e., comparing whether the data from the model and from reality are derived from the same distribution. The two-sample K-S test is calculated as follows:

$$D_{n_M, n_R} = \sup_x |F_{M, n_M}(x) - F_{R, n_R}(x)| \tag{12}$$

where F denotes the empirical distribution of each dataset, which is calculated as follows:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{[-\infty, x]}(X_i) \tag{13}$$

where

$$I_{[-\infty, x]}(X_i) = \begin{cases} 1, & \text{if } X_i \leq x \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

Finally, the null hypothesis is rejected for a given α level if:

$$D_{n_M, n_R} > C(\alpha) \sqrt{\frac{n_M + n_R}{n_M * n_R}} \tag{15}$$

where $c(\alpha)$ is given in the Kolmogorov-Smirnov table.

Chi-square Test: The chi-square (χ^2) test is also a goodness-of-fit test which, similarly to the K-S test, it can also be a one- or two-sample test. The idea behind a two-sample chi-square test, which is more commonly used in model V&V, is that the simulation and operational data are partitioned in i bins, and then the number of points in each bin is observed on whether it is similar on both datasets [46]. Accepting the null hypothesis (H_0) means that the samples are drawn from the same distribution. The chi-square test can be calculated as follows:

$$\chi^2 = \sum_{i=1}^k \frac{(K_M x_{Mi} - K_R x_{Ri})^2}{x_{Mi} + x_{Ri}} \tag{16}$$

which follows the chi-squared distribution, and where i is the number of bins, x_{Mi} and x_{Ri} the observed values from the model and reality respectively, and K_M and K_R constants adjusting the inequality of the observations of the two datasets, which are calculated as follows:

$$K_M = \sqrt{\frac{\sum_{i=1}^k x_{Ri}}{\sum_{i=1}^k x_{Mi}}} \tag{17}$$

$$K_R = \sqrt{\frac{\sum_{i=1}^k x_{Mi}}{\sum_{i=1}^k x_{Ri}}} \tag{18}$$

Anderson-Darling Test: The Anderson-Darling test belongs to the class of quadratic empirical distribution function (EDF)

statistics, which determine whether a sample is drawn from a specific distribution (one-sample) or whether two samples are drawn from the same distribution (two-sample) [47]. The two-sample formula of the test is calculated as follows [48]:

$$AD = \frac{1}{n_M n_R} \sum_{i=1}^{n_M+n_R} (N_i Z_{(n_M+n_R-n_M i)})^2 \frac{1}{i Z_{n_M+n_R-i}} \quad (19)$$

where $Z_{n_M+n_R}$ is the combined and ordered samples of the model and reality and N_i the number of observations in the model that are equal to or smaller than the i^{th} observation in $Z_{n_M+n_R}$.

Cramér–von Mises Criterion: The Cramér–von Mises criterion also belongs to the class of quadratic EDF statistics and is quite similar to the Anderson–Darling test [49]. Compared to the Cramér–von Mises criterion, the Anderson–Darling test places more weight on observations in the tails of the distribution. The two-sample Cramér–von Mises criterion is calculated as follows:

$$T = \frac{U}{n_M n_R (n_M + n_R)} - \frac{4n_M n_R - 1}{6(n_M + n_R)} \quad (20)$$

where

$$U = n_M \sum_{i=1}^{n_M} (r_i - i)^2 + n_R \sum_{j=1}^{n_R} (s_j - j)^2 \quad (21)$$

and $(r_1, r_2, \dots, r_{n_M})$ and $(s_1, s_2, \dots, s_{n_R})$ the ranks of the sorted samples of the model and reality respectively.

Kuiper’s Test: Kuiper’s test is a goodness-of-fit test similar to the Kolmogorov–Smirnov test (K–S test) in the sense that it compares two cumulative distribution functions. Compared to the K–S test, Kuiper’s test is sensitive not only to the median but also to the tail. Compared to the Anderson–Darling test, which also provides equal sensitivity at the tails and at the median, Kuiper’s test is invariant under cyclic transformations of the independent variable [50]. Kuiper’s test is calculated as follows:

$$V = D_+ + D_- \quad (22)$$

where

$$D_+ = \max_{-\infty < x < \infty} [S_M(x) - S_R(x)] \quad (23)$$

$$D_- = \max_{-\infty < x < \infty} [S_R(x) - S_M(x)] \quad (24)$$

$$S_M(x_i) = \frac{i - n_M}{n_M} \quad (25)$$

$$S_R(x_i) = \frac{i - n_R}{n_R} \quad (26)$$

Coefficient of Determination (R^2): R^2 is yet another goodness-of-fit test that indicates the proportion of the variance of the dependent variable that is predicted from the independent variable or variables. The most commonly used extension of R^2 is the adjusted R^2 (\bar{R}^2), which adjusts for the number of explanatory terms in a model relative to the number of data points [51]. \bar{R}^2 is calculated as follows:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n_M - 1}{n_M - k - 1} \quad (27)$$

where

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} \quad (28)$$

$$SS_{residual} = \sum_{i=1}^{n_M} e_i^2 \quad (29)$$

$$SS_{total} = \sum_{i=1}^{n_M} (y_i - \bar{y})^2 \quad (30)$$

and k is the number of independent variables. The closer \bar{R}^2 is to one, the better the model is considered, since the results are explained in a large degree from the variation of the dependent variables and not from the residuals.

Mann–Whitney–Wilcoxon Test: The Mann–Whitney–Wilcoxon (MWW) test, also known as Mann–Whitney U test, is a non-parametric test that tests whether two samples derive from populations having the same distribution [52]. The MWW test can be calculated by first sorting all values from both datasets in an ascending order and assigning numeric ranks starting with 1 from the end of this sorted list. Then, the MWW values for both datasets are computed as follows:

$$U_M = R_M - \frac{n_M(n_M + 1)}{2} \quad (31)$$

$$U_R = R_R - \frac{n_R(n_R + 1)}{2} \quad (32)$$

where R indicates the sum of the ranks for each dataset. Finally, in order to determine whether the two samples derive from the same population, the minimum value between U_M and U_R is compared with the value from the tables.

White Test: The White test is a test for determining whether the variance of a model is constant, i.e., whether the model is homoscedastic (H_0) [53]. The White test is calculated as follows:

$$\hat{e}_i^2 = \delta_0 + \delta_1 \hat{Y}_i + \delta_2 \hat{Y}_i^2 \quad (33)$$

where Y_i are the predicted dependent variables of the model. Upon calculating δ_0 , δ_1 , and δ_2 , the $R_{\hat{e}_i^2}^2$ can be computed and then the $\chi^2 = n_M R_{\hat{e}_i^2}^2$, which can then be tested with 2 degrees of freedom against the null hypothesis.

Glejser Test: The Glejser test also tests for Heteroscedasticity but instead of using the square of the residuals, it uses their absolute values [54]. The Glejser test is calculated as follows:

$$|e_i| = \gamma_0 + \gamma_1 f(x_i) + u_i \quad (34)$$

in which case the most common values for the $f(x_i)$ are: $f(x_i) = x_i$, $f(x_i) = \sqrt{x_i}$, and $f(x_i) = \frac{1}{x_i}$. The γ_1 of the equation with the highest R^2 is then tested and if it is found statistically significant, the null hypothesis of homoscedasticity is rejected.

Spectral Analysis: Spectral analysis tests whether two time series are equivalent [55]. Spectral analysis is a relatively complex statistical test, especially compared to the tests presented so far, and it is calculated as follows:

$$g_i(f) = \frac{1}{\pi} \left[2 \sum_{p=1}^L k_L(p) C_i(p) \cos(f_i(p)) + C_i(0) \right] \quad (35)$$

where $i = \{M, R\}$. $C_i(p)$ is the autocovariance function

$$C_i(p) = \frac{1}{T-p} \sum_{t=1}^{T-p} (x_t - m)(x_{t+p} - m) \quad (36)$$

$k_L(p)$ is a Bartlett weighting function for which several possibilities exists [56], and

$$\begin{aligned} m &= \text{mean of } X(t) \\ T &= \text{total time period} \\ X_t &= \text{observation at time } t \\ f &= \text{frequency in cycles per unit of time} \\ L &= \text{number of lags} \\ p &= \text{number of time periods separating correlated observations } (1, 2, \dots, L-1) \end{aligned}$$

Finally, in order to determine whether the two time series are equivalent, i.e., not rejecting the null hypothesis, the ratio $g_M(f)/g_R(f)$ should satisfy the inequality:

$$e^{-\phi} \leq \frac{g_M(f)}{g_R(f)} \leq e^{\phi} \quad (37)$$

where

$$\phi = Z_{\alpha/2}(4L/3T)^{1/2} \quad (38)$$

and $Z_{\alpha/2}$ = the two tail critical value for the standard normal distribution at a significance level of α .

Durbin-Watson Statistic: The Durbin-Watson statistic tests for the existence of autocorrelation in the residuals from a regression analysis [57]. The statistic is calculated as follows:

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=2}^T e_t^2} \quad (39)$$

where T is the number of observations. The value d is compared to the lower and upper critical values ($d_{L,\alpha}$ and $d_{U,\alpha}$) to test for positive or negative autocorrelation.

The statistical techniques described above as just a sample of the available techniques for simulation model V&V. Nevertheless, it is a representative sample that can be used in the majority of the cases. The aim of this section is to illustrate the various statistical techniques, which facilitates the categorization of these techniques and thus the selection of the most suitable ones given the problem at hand.

B. Step 2: Simulations' and systems' properties and characteristics

This step aims at identifying the properties and characteristics of simulations and the real-world system (hereinafter referred to as system) under study that can potentially influence the selection of methods and techniques.

1) *Simulations' properties:* Since simulations differ from one another in various ways, distinctions are made on whether they represent an existing system, or whether they simulate a system at a microscopic or macroscopic level, or whether they are intended for learning or decision making, and so forth. This is an indication that simulations can be characterized by various properties. Based on literature, this study has identified 10 properties of simulations. The rationale behind selecting those properties was to describe simulations with as much detail as possible. Hence, the properties span on multiple levels. Not all identified properties necessarily influence the selection of V&V methods, therefore, this step is not only about identifying the properties themselves but also determining which are the

ones that really influence the effectiveness of a method; in other words, this step serves as the rationale for choosing those properties of simulations that are applicable to specific V&V methods, and provides for the reasons behind this selection.

The 10 identified properties of simulations are the following:

- 1) Access to the source code of the simulation. Accessibility, or lack of it, influences the selection of a V&V method [58], since several methods require some sort of a check on the code level. Hence, this property is included in the analysis.
- 2) The simulation represents an existing real-system for which real data exist [59]. The existence of, or more importantly the lack of, real data heavily influences the selection process since several methods require real data and thus cannot be used when no real data is available. Hence, this property is also included in the analysis.
- 3) The formalism the simulation is based on, like Discrete Event System Specification (DEVS), Differential Equation Specified System (DESS), System Dynamics, etc. [60]. Several frameworks and methods have been proposed on how to verify and validate DEVS [61][62], DESS [63][64], or system dynamics models [65][66], but they are either application specific or the same method can be used in more than one formalisms, making it independent of the actual formalism. Therefore, while formalisms are an important aspect of simulation modeling, their influence on the V&V method selection is minimal, ergo excluded from the analysis.
- 4) The simulation's worldviews: i) Process Interaction/Locality of Object, ii) Event Scheduling/Locality of Time, iii) Activity Scanning/Locality of State [67]. While worldviews allow for more concise model descriptions by allowing a model specifier to take advantage of contextual information, there is not any evidence from a literature point of view that they have an influence on the V&V method selection, hence, they are excluded from the analysis.
- 5) The fidelity level of the simulation (Low, Medium, High) [68]. While from a literature point of view there is no evidence to support the influence of the level of fidelity on the V&V method selection, common sense dictates that there must be some. Indeed, in order to characterize a simulation as of high fidelity, it must imitate an existing system and real-world data must exist, thus making the comparison and the final characterization possible. Therefore, as discussed in the second property and shown in Table I, the existence of data of the real system influences the V&V method selection, as does the level of fidelity. Yet, since the correlation between real data and high fidelity is almost 1-to-1, the fidelity level is excluded from the analysis for reasons of simplification.
- 6) The type of the simulation (Constructive, Virtual, Live) [69]. This classification, which is adopted by the U.S. Department of Defense [17], should be seen more as a continuum rather than as a discrete characterization. Once a simulation moves towards the Virtual or the Live side of the continuum, it can

also be referred to as 'game'. A game has the distinct characteristic that the game session is succeeded by debriefing, whereby the participants reflect upon the game session to link the content presented during the session with reality [70]. It has been demonstrated that debriefing can in general facilitate validation [71][72]. Moreover, while all methods identified in this paper are suitable for pure simulations (constructive), not all of them are appropriate for games. It would be interesting to examine which of the methods can also be used for validating games. Hence, this property is included in the analysis.

- 7) The purpose the simulation was built for (learning, decision making, etc.). Several case studies on V&V of simulations for different purposes have been reported; in training [73][74], in decision making [75], in concept testing [76], etc., but there are no reports of specific V&V methods being more effective for a certain purpose. Hence, this property is excluded from the analysis.
- 8) The simulation imitates a strictly technical, a socio technical system (STS), or a complex adaptive system (CAS) with multiple agents. There are several studies on modeling and validating simulations for STS [77] and CAS with multiple agents [78][79] but there are no indications that certain V&V methods are more effective for an STS or a CAS. Therefore, this property is excluded from the analysis.
- 9) The application domain of the simulation (logistics, business, physics, etc.). Although the application domain of the simulation plays a significant role in the modeling process, since different approaches are required (Newtonian physics for object movement, Navier–Stokes equations for fluid behavior, etc.) for modeling different systems [80], literature, or more precisely the lack of it, suggests that the V&V process and thus the V&V method selection is not affected by the application domain. Hence, this property is excluded from the analysis.
- 10) The functional (hard goals) and non-functional (soft goals) requirements of the simulation [81]. Validating the simulation's requirements is indeed an important part of the V&V process [82], since validation is always relative to the intended use [83], in other words the use defined in the requirements. Hence, making a distinction between the hard and soft goals is paramount and as such this property is included in the analysis.

2) *Simulations' and systems' characteristics:* Simulations and the systems they imitate can produce a variety of data, which can be characterized in various ways. Moreover, depending on the type of data and on the purpose of the V&V study, different statistical tests are usually necessary, which in turn depend on the produced output. Based on the literature review on the techniques presented in Section II-A2, the characteristics of simulations and systems that influence the selection of techniques are the following:

- 1) Number of datasets. The most usual case in simulation model validation is to have two datasets (model and reality). Nevertheless, there are cases where the number of datasets can be either one, e.g., when

testing whether the model derives from a known distribution like the normal or gamma distributions, or more than two, e.g., when testing the results of more than one models against the operational data.

- 2) Number of variables. The most usual case in simulation model validation is to test one variable, e.g. in railway simulations, this variable is usually the amount of delay. Nevertheless, there are cases where the number of testing variables is more than one, e.g. simultaneously testing longitude and latitude values between model and reality.
- 3) Purpose of the statistical technique. A statistical technique can test for equality of means, the extent to which the data from the model and reality are similarly distributed, the extent to which two time series are equivalent, or it can be used to reduce the model's complexity.
- 4) Known parameters. Statistical techniques are divided in two major categories: parametric and non-parametric. Parametric techniques are the ones that require the mean and variance (μ, σ^2) to be known, whereas non-parametric techniques can deal with cases where these parameters are not known.
- 5) Type of data. The type of data simulations and systems produce range from strictly quantitative to purely qualitative. Usually, statistical techniques suitable for a V&V study should be able to deal with data that are either numerical or categorical (binary).
- 6) Size of samples. Simulation and system data are almost impossible to be normally distributed. Nevertheless, due to the Central Limit Theorem [84], when the size of a sample exceeds 30 (or 40 depending how close to be normally distributed the data are), it is assumed that it follows the normal distribution thus the techniques that work for the normal distribution are applicable.

C. Step 3: Phases of a simulation study

According to Sargent [85], there are 4 distinct phases of V&V: *Data Validation*, *Conceptual Model Validation*, *Model Verification*, and *Operational Validation*. *Data Validation* is concerned with the accuracy of the raw data, as well as the accuracy of any transformation performed on this data. *Conceptual Model Validation* determines whether the theories and assumptions underlying the conceptual model are correct, and whether the model's structure, logic, and mathematical and causal relationships are "reasonable" for the intended purpose of the model. *Model Verification* ensures that the implementation of the conceptual model is correct. Finally, *Operational Validation* is concerned with determining that the model behaves accurately based on its intended purpose. This study adopts Sargent's [85] characterization and aims at using it to classify the methods, in addition to the simulations' properties.

D. Conclusion of the Literature Review

It is evident that selecting one method or technique over another for a V&V study depends on several characteristics of the simulation, the system, the methods, and the techniques, as well as the phase of the simulation study. In Section III, a methodology that combines all three steps aiming at the development of a framework for method and technique selection is proposed.

III. METHODOLOGY

In this section, a methodology for selecting the most appropriate V&V methods (Section III-A) and statistical techniques (Section III-B) for a V&V study is proposed.

A. V&V method selection methodology

As discussed in Section II-B1, dimensions 3, 4, 5, 7, 8, and 9 are perceived to have little influence on the method selection, hence, there are excluded from the analysis. On the other hand, the purpose of the method selection, discussed in Section II-C, seems to be crucial; in other words, it is important to differentiate on whether the selected method will be used for data validation, conceptual model validation, model verification, or operational validation. Therefore, the list of the dimensions is refined, and is expressed in questions, as follows:

- 1) Does the V&V method require access to the simulation model's source code?

Possible answers: Yes or No. A positive answer to this question means that this method can only be used when the person or persons performing the V&V have access to the simulation's source code, whereas a negative answer means that it can be used in any occasion regardless of the accessibility to the simulation model's source code. It should be noted that the current study - and consequently this dimension - is not concerned with the specific programming language the simulation is built on (Assembly, C++, NetLogo, etc.), but solely with whether the application of a V&V method depends upon having access to the source code.

- 2) Does the V&V method require data from the real system?

Possible answers: Yes or No. A positive answer to this question means that this method can only be used when data from the real system are available, whereas a negative answer means that it can be used in any occasion regardless of the availability of data from the real system. It should be noted that the current study - and consequently this dimension - is not concerned with the nature of the data in general (qualitative or quantitative), but solely with their existence and availability.

- 3) Is the V&V method suitable for a game V&V study?

Possible answers: Yes or No. While all methods are suitable for pure simulations, some of them will be also suitable for games in particular. Although games often have a simulation model running on the background, in which case all methods would be applicable, in this study the term *game* is used to describe the layer that is on top of the simulation model and refers to the players' interaction.

- 4) For what type of requirements is the V&V method more suitable?

Possible answers: Hard (Functional), or Soft (Non-Functional), or Both. A method might be focused on either the functional part or the non-functional part of the model or on both.

- 5) For which type of study is the V&V method more suitable?

Possible answers: Data Validation (D. Val.), Conceptual Model Validation (C.M. Val.), Model Verification (M. Ver.), or Operational Validation (O. Val.). A

method might be suitable for one or more of the available categories.

Table I summarizes the results of the analysis. The intended use of Table I is to act as a filtering mechanism. Whenever an individual or a team wants to verify and/or validate a simulation model, they can utilize this table to narrow down the applicable methods according to the different properties of the simulation at hand. The selection process is shown in Figure 1.

With regards to the first property, i.e., the accessibility to the source code, and in contrary to the second property, access to the source code does not imply that the methods categorized under "Yes" are stronger. Usually, access to the source code is associated with verification and in some cases conceptual model validation.

With regards to the second property, i.e., the availability of data from the real system, the methods categorized under "No" can be used irrespective of whether real data exist. Nevertheless, the methods categorized under "Yes" are more powerful in the sense that, if used appropriately, they provide evidence or a data trace of how the simulation should work. Hence, whenever real data are available, the methods categorized under "Yes" should be preferred, unless an alternative method is definitely more suitable.

With regards to the third property, i.e., the suitability of certain methods for the V&V of games, informal methods [11] seem to be the ones suitable for games. This is a preliminary conclusion that is expected to an extent. In games representing Complex Adaptive Systems (CAS), experts' opinion plays an important, and perhaps the most important, role [86], regardless of the game's level of fidelity [87] or use of technology [88]. It should be noted that although the term *Games* assumes both high-tech (computer-based) and low-tech (e.g., tabletop) games, the selection in Table I was made with a bias towards the high-tech games.

With regards to the fourth and fifth property, i.e., the type of requirements being tested and the purpose of the V&V study respectively, the answers are more or less self-explanatory. Some methods are more suitable for testing one type of requirement. As an example, regression testing is more appropriate for functional requirements (hard goals). Other V&V methods are better suited for one purpose, such as Structural (White-box) testing, which is more appropriate for conceptual model validation, while others are more suitable for testing both types of requirements (e.g., Graphical comparisons), or for more than one purpose (e.g., Trace-Driven Input Testing).

The novelty of the proposed framework does not lie in the content of Table I per se, but on the idea that the list of methods can be narrowed down to a manageable level, thus making the V&V of a simulation better grounded, faster, more accurate, and more cost effective.

There is a threat towards the validity of the content on Table I. The line between whether data from the real system are needed, or whether access to the source code is needed, or whether a specific requirement is definitely functional or non-functional, or whether the purpose is to validate the data, the conceptual model, the operational ability of the model, or to just verify the model, is not always clear and well defined. In Section V, future steps are proposed aiming at addressing and mitigating the above mentioned threat.

TABLE I. LIST OF V&V METHODS & PROPERTIES OF SIMULATIONS.

Method	Source Code	Real Data	Game	Requirements	Purpose
Acceptance Testing	No	No	Yes	Both	O. Val.
Alpha Testing	No	No	Yes	Both	O. Val.
Assertion Checking	Yes	No	No	Hard	M. Ver.
Audit	Yes	No	Yes	Soft	M. Ver.
Beta Testing	No	No	Yes	Both	O. Val.
Bottom-Up Testing	Yes	No	No	Both	M. Ver.
Cause-Effect Graphing	Yes	No	No	Hard	M. Ver.
Comparison Testing	No	No	No	Both	C.M. Val.
Compliance Testing: → Authorization Testing	No	No	No	Soft	M. Ver.
→ Performance Testing	No	No	No	Soft	M. Ver.
→ Security Testing	No	No	No	Soft	M. Ver.
→ Standards Testing	No	No	No	Soft	M. Ver.
Control Analysis: → Calling Structure Analysis	Yes	No	No	Hard	C.M. Val.
→ Concurrent Process Analysis	Yes	No	No	Hard	M. Ver.
→ Control Flow Analysis	Yes	No	No	Hard	C.M. Val.
→ State Transition Analysis	Yes	No	No	Hard	D. Val. & M. Ver.
Data Analysis: → Data Dependency Analysis	Yes	No	No	Hard	D. Val. & M. Ver.
→ Data Flow Analysis	Yes	No	No	Hard	D. Val. & M. Ver.
Debugging	Yes	No	No	Both	M. Ver.
Desk Checking	Yes	No	Yes	Both	M. Ver.
Documentation Checking	Yes	No	Yes	Both	C.M. Val.
Execution Testing: → Execution Monitoring	No	No	No	Hard	C.M. Val.
→ Execution Profiling	No	No	No	Hard	C.M. Val.
→ Execution Tracing	Yes	No	No	Hard	C.M. Val.
Face Validation	No	Yes	Yes	Both	O. Val.
Fault/Failure Analysis	No	No	No	Hard	C.M. Val.
Fault/Failure Insertion Testing	No	No	No	Hard	C.M. Val.
Field Testing	No	Yes	No	Both	O. Val.
Functional (Black-Box) Testing	No	Yes	No	Hard	C.M. Val.
Graphical Comparisons	No	Yes	Yes	Both	O. Val.
Induction	No	No	No	Both	C.M. Val.
Inference	No	No	No	Both	C.M. Val.
Inspections	No	No	No	Both	C.M. Val.
Interface Analysis: → Model Interface Analysis	No	No	No	Soft	C.M. Val.
→ User Interface Analysis	No	No	Yes	Soft	O. Val.
Interface Testing: → Data Interface Testing	No	No	No	Soft	D. Val.
→ Model Interface Testing	No	No	No	Soft	C.M. Val.
→ User Interface Testing	No	No	Yes	Soft	O. Val.
Lambda Calculus	Yes	No	No	Hard	M. Ver.
Logical Deduction	No	No	No	Both	All
Object-Flow Testing	No	No	No	Hard	O. Val.
Partition Testing	Yes	No	No	Hard	C.M. Val.
Predicate Calculus	Yes	No	No	Hard	M. Ver.
Predicate Transformations	No	Yes	No	Hard	M. Ver.
Predictive Validation	No	Yes	No	Hard	O. Val.
Product Testing	No	No	Yes	Both	O. Val.
Proof of Correctness	Yes	No	No	Hard	C.M. Val. & M. Ver.
Regression Testing	Yes	No	No	Hard	M. Ver.
Reviews	No	No	Yes	Both	C.M. Val.
Semantic Analysis	Yes	No	No	Both	M. Ver.
Sensitivity Analysis	No	No	No	Hard	O. Val.
Special input testing: → Boundary Value Testing	Yes	No	No	Both	M. Ver.
→ Equivalence Partitioning Testing	No	No	No	Hard	O. Val.
→ Extreme Input Testing	No	No	No	Hard	O. Val.
→ Invalid Input Testing	No	No	No	Hard	O. Val.
→ Real-Time Input Testing	No	Yes	No	Hard	O. Val.
→ Self-Driven Input Testing	No	No	No	Hard	O. Val.
→ Stress Testing	No	No	No	Hard	O. Val.
→ Trace-Driven Input Testing	Yes	Yes	No	Both	D. Val. & C.M. Val.
Structural (White-box) Testing: → Branch Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
→ Condition Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
→ Data Flow Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
→ Loop Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
→ Path Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
→ Statement Testing	Yes	No	No	Both	C.M. Val. & M. Ver.
Structural Analysis	No	No	No	Hard	C.M. Val.
Submodel/Module Testing	No	No	No	Both	C.M. Val.
Symbolic Debugging	Yes	No	No	Hard	M. Ver.
Symbolic Evaluation	Yes	No	No	Hard	C.M. Val.
Syntax Analysis	Yes	No	No	Hard	M. Ver.
Top-Down Testing	Yes	No	No	Both	C.M. Val.
Traceability Assessment	Yes	Yes	No	Both	C.M. Val.
Turing Test	No	Yes	No	Both	O. Val.
Visualization/Animation	No	Yes	Yes	Both	O. Val.
Walkthroughs	No	No	Yes	Both	C.M. Val.

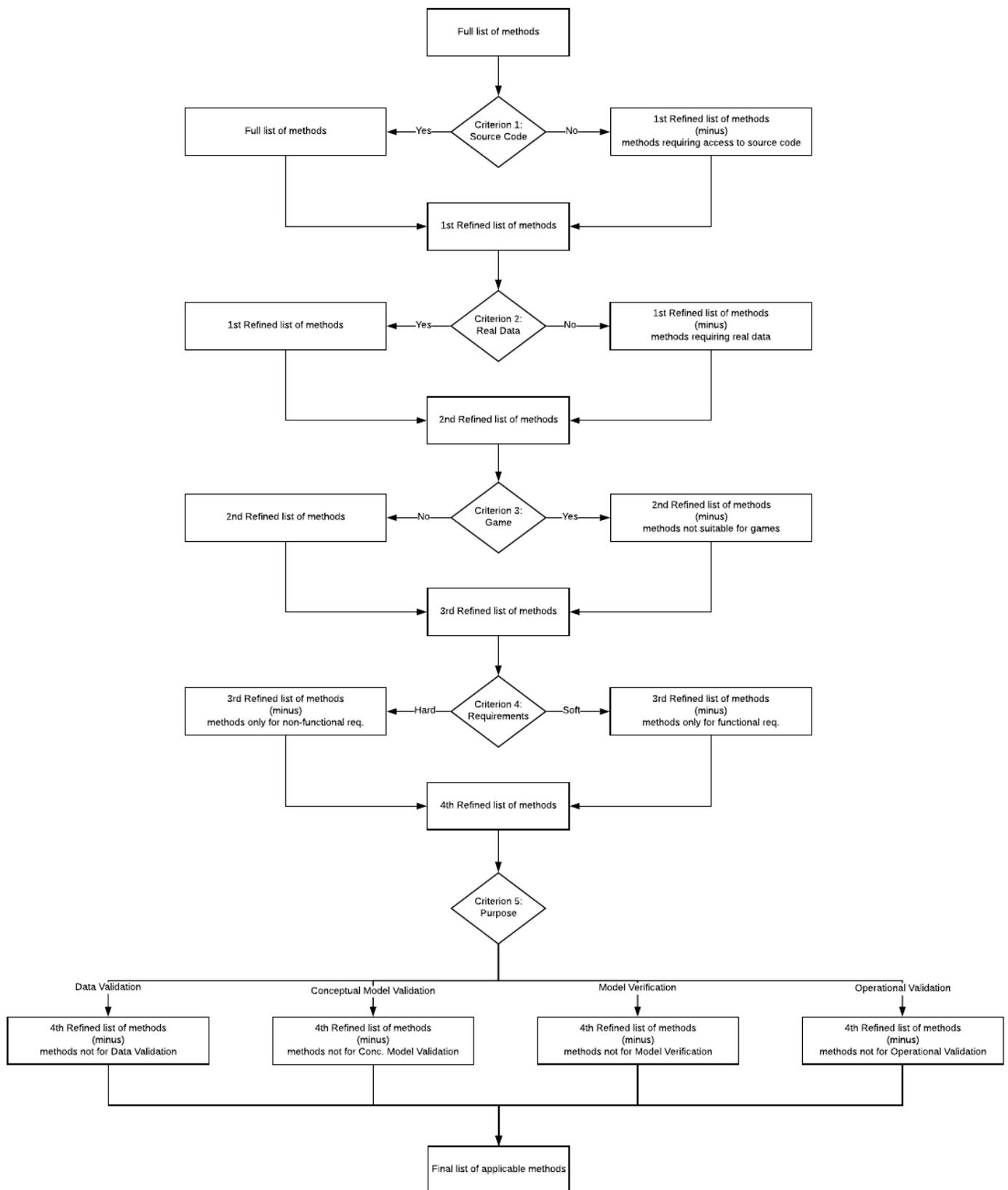


Figure 1. The flow diagram of the selection process of methods.

TABLE II. LIST OF STATISTICAL TECHNIQUES.

Statistical Techniques	# of datasets	# of variables	Purpose	known parameters	Type of data	Sample size
t-Test	1 or 2	1	Mean equality	Yes	Numerical	Any
Hottelling's T^2 Test	1 or 2	>1	Mean equality	Yes	Numerical	Any
Analysis of Variance	>2	1	Mean equality	Yes	Numerical	Any
Multivariate Analysis of Variance	>2	>1	Mean equality	Yes	Numerical	Any
Simultaneous Confidence Intervals	1 or 2	>1	Mean equality	Yes	Numerical	Any
Factor Analysis	1	>1	Complexity reduction	Yes	Numerical	Any
Principal Component Analysis	1	>1	Complexity reduction	Yes	Numerical	Any
Kolmogorov-Smirnov Test	1 or 2	1	Goodness-of-fit	No	Numerical	Any
Chi-squared Test	1 or 2	1	Goodness-of-fit	No	Numerical & Categorical	Any
Anderson-Darling Test	1 or 2	1	Goodness-of-fit	No	Numerical	Any
Cramér-von Mises Criterion	1 or 2	1	Goodness-of-fit	No	Numerical	Any
Kuiper's Test	1 or 2	1	Goodness-of-fit	No	Numerical	Any
Coefficient of Determination	2	1	Goodness-of-fit	Yes	Numerical	Any
Mann-Whitney-Wilcoxon Test	2	1	Mean equality	No	Numerical & Categorical	Small
White Test	2	1	Heteroscedasticity	Yes	Numerical	Any
Glejser Test	2	1	Heteroscedasticity	Yes	Numerical	Any
Spectral Analysis	2	1	Time Series analysis	Yes	Numerical	Any
Durbin-Watson Statistic	2	1	Time Series analysis	Yes	Numerical	Any

B. Statistical technique selection methodology

The list of simulations' and systems' characteristics that influence the selection of techniques, which are explained in more detail in Section II-B2, are expressed in questions, as follows:

- 1) How many different datasets are going to be examined?
Possible answers: 1, 2, and/or >2.
- 2) How many different variables are going to be examined?
Possible answers: 1 or >1.
- 3) What is the purpose of the statistical test?
Possible answers: Mean equality, Complexity reduction, Goodness-of-fit, Heteroscedasticity, or Time Series analysis.
- 4) Are the sample parameters (μ, σ^2) known?
Possible answers: Yes or No.
- 5) What kind of data are going to be examined?
Possible answers: Numerical or Categorical.
- 6) What is the sample size?
Possible answers: Large, Small, or Any.

Table II summarizes the results of the analysis. Similarly to Table I, the intended use of Table II is to act as a filtering mechanism. Whenever an individual or a team wants to verify and/or validate a simulation model, they can utilize this table to narrow down the applicable techniques according to the different characteristics of the simulation at hand and the system. It should be noted that another significant factor in selecting a technique is the statistical power of the technique, i.e., the probability that the null hypothesis (H_0) is correctly rejected for the alternative hypothesis (H_1). The statistical power of a technique is not predetermined, which is the reason it is not included in this analysis. Nevertheless, the Neyman-Pearson lemma [89] is a test that determines which technique is the one with the greatest statistical power given several attributes, like the sample size and the statistical significance.

IV. A CASE STUDY

In this section, a case study illustrates how the framework, through the use of Table I and Table II, can be used. The case study is a computer simulation of a particular instantiation of the Dutch railways. The authors were assigned to validate the simulation model with regards to punctuality, in other words the precision of the delays of trains in the model.

In more detail, the simulation model was built on the Friso simulation package [90]. FRISO is ProRail's, the Dutch infrastructure manager, in-house simulation environment. Being a microscopic simulation environment, FRISO has the potential to - and depending on the model it usually does so- simulate the railway network in a detailed manner; it has the ability to depict the network down to a switch level, which is the case with this model. The model was built in 2014 and it simulates the train operations in one of the most heavily utilized sections (Amsterdam Central station - Utrecht Central station) of one of the largest corridors in the Netherlands (A2), during the whole month of June 2013. The intended use of the model was to examine the punctuality of the timetable with the particular focus being the Amsterdam and Utrecht central stations. A more in depth description of the model, including its input, output, and the final results can be found in [91].

With regards to the methods, the initial list, as it is shown in Table I, consists of 75 methods. Then with every step, the list is narrowed down. For this particular study, the selection process for each property, as shown in Figure 2, was as follows:

- 1) Access to the source code was not available; *Answer: No.* Using this criterion the available methods were reduced to 42.
- 2) There were available data from the real system; *Answer: Yes.* Using this criterion eliminated 33 more methods returning a total of 9 available methods. Nonetheless, all 42 methods could have been used in this particular case.
- 3) The main focus was on the punctuality, ergo functional (hard) requirements, but comments were also expected on the non-functional (soft) requirements; *Answer: Both (but main focus on hard).* If on the previous criterion *Yes* was chosen as an option, choosing either *Both* or *Hard* on this criterion would leave the list intact (Total 9 methods).
- 4) The study was mainly concerned with the operational validity of the simulation, but to a degree also with the conceptual model validity; *Answer: C.M. Val & O. Val.* Using this criterion and based on the selections on the previous criteria, the final number of available methods was reduced to 1 for the conceptual model validation and 7 for the operational validation.

For the operational validation, which was the primary

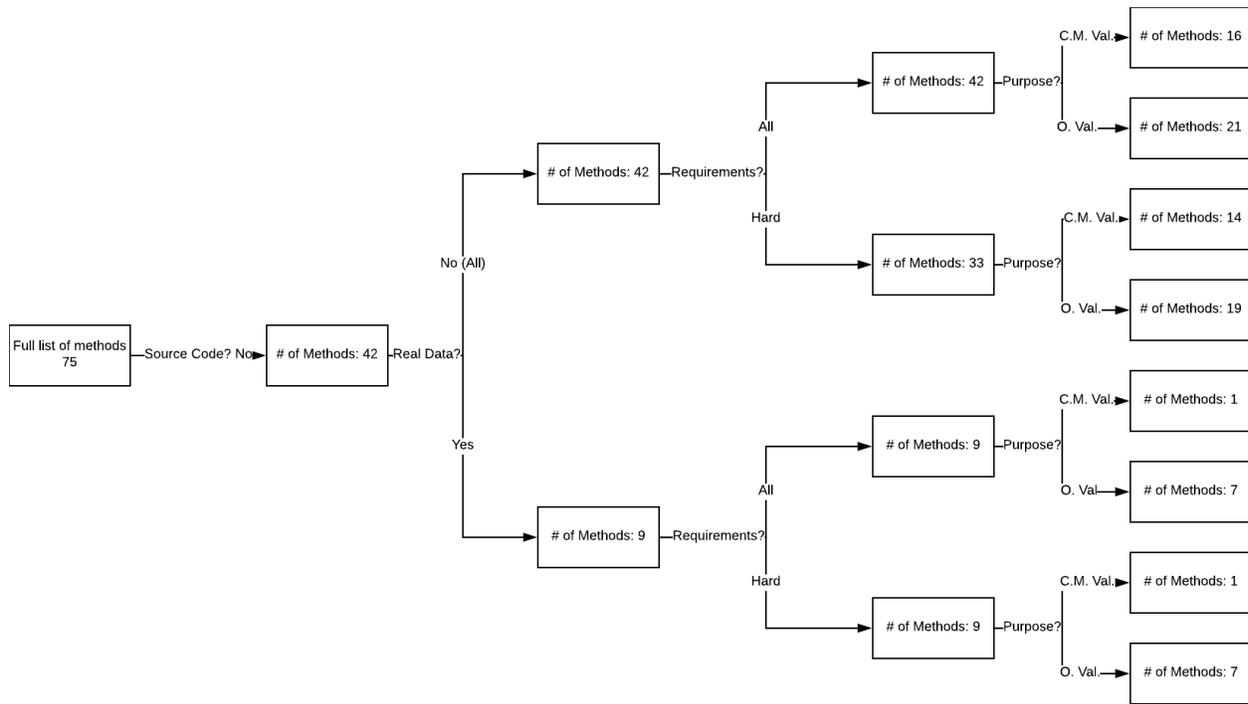


Figure 2. A tree graph of the method selection process.

interest of the study, the final list of the seven methods is shown in Table III. Out of this list, in total four methods were used, namely: the *Face Validation*, *Graphical Comparisons*, *Predictive Validation*, and *Turing Test*. *Predictive Validation* was first used to handle the initial datasets (simulation dataset & operational dataset) and to produce results for the different statistical tests. Then, a combination of the remaining three methods was used to ascertain the validity of the simulation.

With regards to the techniques, the initial list, as shown in Table II, consists of 18 techniques. Then with every step, the list is narrowed down. For this particular study, the selection process for each characteristic was as follows:

- 1) The model's and reality's output were examined; *Answer: 1 or 2*. Using this criteria reduced the available techniques to 16.
- 2) The amount of delays was the focus; *Answer: 1*. Using this criterion eliminated 4 more techniques totaling in 12 available techniques.
- 3) The study was concerned with whether the delays between the model and reality were similarly distributed and whether the averages were significantly different; *Answer: Mean equality and Goodness-of-fit*. Using this criterion resulted in 2 suitable techniques for mean equality and 6 for goodness-of-fit.
- 4) The sample parameters (μ, σ^2) were known; *Answer: Yes*. This is a criterion that only influences the results if the answer is *No*, since the non-parametric techniques can still be used when the mean and variance are known. Therefore the number of techniques remained the same.
- 5) The delays were in seconds, hence numerical; *Answer: Numerical*. Using this criterion eliminated 1 techniques for the mean equality, resulting in 1

- 6) technique for mean equality and 6 for goodness-of-fit. *Answer: Large*. This last criterion did not further reduced the number of techniques, since the only techniques suitable for small datasets (*Mann-Whitney-Wilcoxon Test*) had been eliminated in a previous step.

For testing the equality of means, the only suitable techniques, i.e., t-test, was used. Whereas for testing the goodness-of-fit, from the 6 suitable techniques, the Kolmogorov-Smirnov and chi-squared test were used.

TABLE III. REFINED LIST OF V&V METHODS OF THE CASE STUDY.

Method	1	2	3	4
Face Validation	No	Yes	Both	O. Val.
Field Testing	No	Yes	Both	O. Val.
Graphical Comparisons	No	Yes	Both	O. Val.
Predictive Validation	No	Yes	Hard	O. Val.
Real-Time Input Testing	No	Yes	Hard	O. Val.
Turing Test	No	Yes	Both	O. Val.
Visualization/Animation	No	Yes	Both	O. Val.

In this section, the use of the proposed framework demonstrates clearly its effectiveness. As shown in Table III, the initial list of 75 methods was narrowed down in a matter of minutes to the manageable level of seven; and similar reduction occurred in the techniques. By all means, the effectiveness of the framework is not only evident due to its time-saving nature but also due to the fact that it ensures that the chosen methods and techniques are appropriate for the simulation and the system at hand as well as for the purpose of the V&V study.

V. CONCLUSION & FUTURE WORK

In this paper, a framework for simulation validation and verification method and statistical technique selection was proposed. Various properties and characteristics of simulations

and systems were taken into account and it was shown that indeed some of these influence the method and technique selection and thus, the final results of the simulation study.

Moreover, the framework was applied on a case study, as a first step towards verifying its effectiveness. The case study showed that the framework is an effective time-saving tool, which also provides a safety net for choosing the methods and techniques that best serve the intended purpose of the simulation and the V&V study.

With regards to future work, additional simulation properties should be identified that may potentially influence the method selection, or some of the discarded properties, identified in Section II-B, might prove to be more influential than initially acknowledged. Moreover, there is a need to further verify the connection of each method to the simulation model's properties and the purpose for which they are more suitable; in other words, it should be verified that the answers on columns 2-6 in Table I are correct. With regards to the techniques, a more extensive list analyzed in the same way as in Section III-B would provide for an improved guide towards selecting the most effective techniques given the problem at hand. Finally, more case studies, from the authors and more importantly from researchers unrelated to the authors, both in pure simulations and in games, would further strengthen the validity and applicability of the framework.

Nevertheless, this paper paves the way for future research in the topic, and as discussed earlier, the main contribution of the framework does not lie in the results presented in Table I and Table II, but is related to the identification of the relationships between the methods, the techniques, the simulation's and system's properties, and the purpose of the V&V study. Therefore, it is of utmost importance that any future research be focused on these relationships.

ACKNOWLEDGMENT

This research is funded through the Railway Gaming Suite 2 program, a joint project by ProRail and Delft University of Technology.

REFERENCES

- [1] B. Rongas, S. Meijer, and A. Verbraeck, "A framework for simulation validation & verification method selection," in *SIMUL 2017: The Ninth International Conference on Advances in System Simulation*, Athens, Greece, 2017, pp. 35–40.
- [2] J. W. Forrester, *World dynamics*. Wright-Allen Press, 1971.
- [3] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens III, *The limits to growth*. New York, U.S.A.: Universe Books, 1972.
- [4] D. L. Meadows, W. W. Behrens III, D. H. Meadows, R. F. Naill, J. Randers, and E. Zahn, *Dynamics of growth in a finite world*. Cambridge, Massachusetts: Wright-Allen Press, 1974.
- [5] M. Janssen and B. De Vries, "Global modelling: Managing uncertainty, complexity and incomplete information," in *Validation of Simulation Models*. SISWO, Amsterdam, The Netherlands, 1999, pp. 45–69.
- [6] W. D. Nordhaus, "World dynamics: Measurement without data," *The Economic Journal*, vol. 83, no. 332, 1973, pp. 1156–1183.
- [7] O. Balci, "Validation, verification, and testing techniques throughout the life cycle of a simulation study," *Annals of Operations Research*, vol. 53, no. 1, 1994, pp. 121–173.
- [8] S. Schlesinger, R. E. Crosbie, R. E. Gagné, G. S. Innis, C. S. Lalwani, J. Loch, R. J. Sylvester, R. D. Wright, N. Kheir, and D. Bartos, "Terminology for model credibility," *Simulation*, vol. 32, no. 3, 1979, pp. 103–104.
- [9] O. Balci, "Verification, validation, and certification of modeling and simulation applications," in *Proceedings of the 35th Conference on Winter Simulation*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, Eds. New Orleans, Louisiana, USA: Winter Simulation Conference, 2003, pp. 150–158.
- [10] R. G. Sargent, "Verification and validation of simulation models," in *Proceedings of the 37th Conference on Winter Simulation*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, Eds. Orlando, Florida, USA: Winter Simulation Conference, 2005, pp. 130–143.
- [11] O. Balci, "Verification, validation, and testing," in *Handbook of Simulation*, J. Banks, Ed. Engineering & Management Press, 1998, ch. 10, pp. 335–393.
- [12] W. E. Perry, *Effective methods for software testing*. Wiley Publishing Inc., 2007.
- [13] B. Beizer, *Software testing techniques (2nd edition)*. Van Nostrand Reinhold Company Limited, 1990.
- [14] L. A. Miller, E. Groundwater, and S. M. Mirsky, "Survey and assessment of conventional software verification and validation methods," in No. NUREG/CR-6018; EPRI-TR-102106; SAIC-91/6660. Nuclear Regulatory Commission, Washington, DC (United States). Div. of Systems Research; Science Applications International Corp., Reston, VA (United States), 1993.
- [15] G. J. Myers, T. Badgett, T. M. Thomas, and C. Sandler, *The art of software testing*. John Wiley & Sons, Inc., 2011.
- [16] R. S. Pressman, *Software engineering: A practitioner's approach (8th edition)*. McGraw-Hill, New York, NY, 2015.
- [17] U.S. Department of Defense, "DoD Modeling and Simulation (M&S) Glossary," Tech. Rep., 1997.
- [18] C. Rattray, *Specification and verification of concurrent systems*. Springer-Verlag London, 1990.
- [19] O. Balci, "The implementation of four conceptual frameworks for simulation modeling in high-level languages," in *Proceedings of the 20th Conference on Winter Simulation*, M. A. Abrams, P. L. Haigh, and J. C. Comfort, Eds. San Diego, California, USA: ACM, 1988, pp. 287–295.
- [20] R. H. Dunn, "The quest for software reliability," in *Handbook of Software Quality Assurance*. New York: Van Nostrand Reinhold, 1987.
- [21] R. W. Adrion, M. A. Branstad, and J. C. Cherniavsky, "Validation, verification, and testing of computer software," *ACM Computing Surveys (CSUR)*, vol. 14, no. 2, 1982, pp. 159–192.
- [22] C. F. Hermann, "Validation problems in games and simulations with special reference to models of international politics," *Behavioral Science*, vol. 12, no. 3, 1967, pp. 216–231.
- [23] R. Shannon and J. D. Johannes, "Systems simulation: The art and science," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 6, no. 10, 1976, pp. 723–724.
- [24] *Systems and software engineering – Vocabulary*. IEEE, 2011.
- [25] M. S. Martis, "Validation of simulation based models: A theoretical outlook," *Electronic Journal of Business Research Methods*, vol. 4, no. 1, 2006, pp. 39–46.
- [26] I. M. Copi, C. Cohen, and D. E. Flage, *Essentials of logic*. Taylor & Francis, 2016.
- [27] S. R. Schach, *Classical and object-oriented software engineering (8th edition)*. McGraw-Hill, 2011.
- [28] H. P. Barendregt, *The Lambda Calculus: Its syntax and semantics*, 1984.
- [29] I. Burnstein, *Practical software testing: A process-oriented approach*. Springer-Verlag New York, 2006.
- [30] E. W. Dijkstra, "Guarded commands, non-determinacy and a calculus for the derivation of programs," in *Language Hierarchies and Interfaces*. Lecture Notes in Computer Science, vol. 46, F. Bauer et al., Ed. Springer, Berlin, Heidelberg, 1976, pp. 111–124.
- [31] J. R. Emshoff and R. L. Sisson, *Design and use of computer simulation models*. MacMillan, New York, 1970.
- [32] R. C. Backhouse, *Program construction and verification*. Prentice-Hall International, 1986.
- [33] R. B. Whitner and O. Balci, "Guidelines for selecting and using simulation model verification techniques," in *Proceedings of the 21st Conference on Winter Simulation*, E. MacNair, K. Musselman, and P. Heidelberger, Eds. Washington, D.C., USA: ACM, 1989, pp. 559–568.

- [34] J. C. King, "Symbolic execution and program testing," *Communications of the ACM*, vol. 19, no. 7, 1976, pp. 385–394.
- [35] I. Sommerville, *Software engineering* (9th edition). Addison-Wesley, Reading, MA, 2004.
- [36] L. W. Schruben, "Establishing the credibility of simulations," *Simulation*, vol. 34, no. 3, 1980, pp. 101–105.
- [37] D. L. Hahs-Vaughn and R. G. Lomax, *An introduction to statistical concepts*. Routledge, 2013.
- [38] B. L. Welch, "The generalization of 'Student's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, 1947, p. 28.
- [39] H. Hotelling, "The generalization of Student's ratio," in *American Mathematical Society*, Berkeley, CA, USA, 1931.
- [40] T. H. Naylor and J. M. Finger, "Verification of computer simulation models," *Management Science*, vol. 14, no. 2, 1967, pp. B-92–B-101.
- [41] R. G. Lomax and D. L. Hahs-Vaughn, *Statistical concepts: A second course*. Taylor & Francis Group, 2013.
- [42] R. T. Warne, "A primer on multivariate analysis of variance (MANOVA) for behavioral scientists," *Practical Assessment, Research & Evaluation*, vol. 19, no. 17, 2014, pp. 1–10.
- [43] O. Balci and R. G. Sargent, "Validation of simulation models via simultaneous confidence intervals," *American Journal of Mathematical and Management Sciences*, vol. 4, no. 3-4, 1984, pp. 375–406.
- [44] I. T. Jolliffe, *Principal Component Analysis and Factor Analysis*. Springer, 1986.
- [45] I. M. Chakravarty, J. D. Roy, and R. G. Laha, *Handbook of methods of applied statistics*. John Wiley & Sons, 1967.
- [46] R. A. Fisher, "The conditions under which χ^2 measures the discrepancy between observation and hypothesis," *Journal of the Royal Statistical Society*, vol. 87, no. 3, 1924, pp. 442–450.
- [47] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes," *The Annals of Mathematical Statistics*, vol. 23, no. 2, 1952, pp. 193–212.
- [48] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests," *The Annals of Mathematical Statistics*, vol. 28, no. 4, 1957, pp. 823–838.
- [49] H. Cramér, "On the composition of elementary errors," *Scandinavian Actuarial Journal*, vol. 1928, no. 1, 1928, pp. 13–74.
- [50] N. H. Kuiper, "Tests concerning random points on a circle," *Indagationes Mathematicae*, vol. 63, 1960, pp. 38–47.
- [51] H. Theil, *Economic forecasts and policy*. Amsterdam. The Netherlands: North-Holland.
- [52] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics*, vol. 18, no. 1, 1947, pp. 50–60.
- [53] H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, vol. 48, no. 4, 1980, pp. 817–838.
- [54] H. Glejser, "A new test for heteroskedasticity," *Journal of the American Statistical Association*, vol. 64, no. 325, 1969, p. 316.
- [55] J. A. Fitzsimmons, "The use of spectral analysis to validate planning models," *Socio-Economic Planning Sciences*, vol. 8, no. 3, 1974, pp. 123–128.
- [56] G. M. Jenkins and D. G. Watts, *Spectral analysis and its applications*. Holden-Day, San Francisco, 1969.
- [57] J. Durbin and G. S. Watson, "Testing for serial correlation in least squares regression: I," *Biometrika*, vol. 37, no. 3/4, 1950, p. 409.
- [58] W. F. van Gunsteren and A. E. Mark, "Validation of molecular dynamics simulation," *The Journal of Chemical Physics*, vol. 108, no. 15, 1998, pp. 6109–6116.
- [59] J. P. Kleijnen, "Verification and validation of simulation models," *European Journal of Operational Research*, vol. 82, no. 1, 1995, pp. 145–162.
- [60] H. Vangheluwe, J. de Lara, and P. J. Mosterman, "An introduction to multi-paradigm modelling and simulation," in *Proceedings of the AIS'2002 Conference (AI, Simulation and Planning in High Autonomy Systems)*, F. Barros and N. Giambiasi, Eds., Lisboa, Portugal, 2002, pp. 9–20.
- [61] J. H. Byun, C. B. Choi, and T. G. Kim, "Verification of the DEVS model implementation using aspect embedded DEVS," in *Proceedings of the 2009 Spring Simulation Multiconference*. San Diego, CA, USA: Society for Computer Simulation International, 2009, p. 151.
- [62] H. Saadawi and G. Wainer, "Verification of real-time DEVS models," in *Proceedings of the 2009 Spring Simulation Multiconference*. San Diego, CA, USA: Society for Computer Simulation International, 2009, p. 143.
- [63] M. D. Di Benedetto, S. Di Gennaro, and A. D'Innocenzo, "Diagnosability verification for hybrid automata," in *Hybrid Systems: Computation and Control*, A. Bemporad, A. Bicchi, and G. Buttazzo, Eds. Springer, Berlin, Heidelberg, 2007, pp. 684–687.
- [64] J. Jo, S. Yoon, J. Yoo, H. Y. Lee, and W.-T. Kim, "Case study: Verification of ECML model using SpaceEx," in *Korea-Japan Joint Workshop on ICT*, 2012, pp. 1–4.
- [65] Y. Barlas, "Model validation in system dynamics," in *Proceedings of the 1994 International System Dynamics Conference*, Sterling, Scotland, 1994, pp. 1–10.
- [66] J. W. Forrester and P. M. Senge, "Tests for building confidence in system dynamics models," in *System Dynamics*, TIMS Studies in Management Sciences, 14, 1980, pp. 209–228.
- [67] M. C. Overstreet and R. E. Nance, "Characterizations and relationships of world views," R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, Eds. Washington Hilton and Towers, Washington, D.C., U.S.A.: ACM, 2004.
- [68] D. Liu, N. D. Macchiarella, and D. A. Vincenzi, "Simulation fidelity," in *Human Factors in Simulation and Training*, 2008.
- [69] J. E. Morrison and L. L. Meliza, "Foundations of the after action review process," Alexandria, VA, p. 82, 1999.
- [70] R. Fanning and D. Gaba, "The role of debriefing in simulation-based learning," *Simulation in Healthcare*, vol. 2, no. 2, 2007, pp. 115–125.
- [71] J. van den Hoogen, J. Lo, and S. Meijer, "Debriefing in gaming simulation for research: Opening the black box of the non-trivial machine to assess validity and reliability," in *Proceedings of the 2014 Winter Simulation Conference*, A. Tolk, S. Y. Djalilo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, Eds. Savannah, Georgia, USA: IEEE Press, 2014, pp. 3505–3516.
- [72] J. Lo, J. van den Hoogen, and S. Meijer, "Using gaming simulation experiments to test railway innovations: Implications for validity," in *Proceedings of the 2013 Winter Simulation Conference*, R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, Eds. Washington, D.C., USA: IEEE Press, 2013, pp. 1766–1777.
- [73] B. Zevin, J. S. Levy, R. M. Satava, and T. P. Grantcharov, "A Consensus-based framework for design, validation, and implementation of simulation-based training curricula in surgery," *Journal of the American College of Surgeons*, vol. 215, no. 4, 2012, pp. 580–586.e3.
- [74] P. J. Morgan, D. Cleave-Hogg, S. DeSousa, and J. Tarshis, "High-fidelity patient simulation: Validation of performance checklists," *British Journal of Anaesthesia*, vol. 92, no. 3, 2004, pp. 388–392.
- [75] S. I. Gass, "Decision-aiding models: Validation, assessment, and related issues for policy analysis," *Operations Research*, vol. 31, no. 4, 1983, pp. 603–631.
- [76] R. R. Nemani and S. W. Running, "Testing a theoretical climate-soil-leaf area hydrologic equilibrium of forests using satellite data and ecosystem simulation," *Agricultural and Forest Meteorology*, vol. 44, no. 3-4, 1989, pp. 245–260.
- [77] A. Mavin and N. Maiden, "Determining socio-technical systems requirements: Experiences with generating and walking through scenarios," in *Proceedings of the 11th IEEE International Conference on Requirements Engineering*. IEEE Comput. Soc, 2003, pp. 213–222.
- [78] F. Nilsson and V. Darley, "On complex adaptive systems and agent-based modelling for improving decision-making in manufacturing and logistics settings," *International Journal of Operations & Production Management*, vol. 26, no. 12, 2006, pp. 1351–1373.
- [79] M. A. Louie and K. M. Carley, "Balancing the criticisms: Validating multi-agent models of social systems," *Simulation Modelling Practice and Theory*, vol. 16, no. 2, 2008, pp. 242–256.
- [80] F. Landriscina, *Simulation and learning*. Springer, 2013.
- [81] J. Mylopoulos, L. Chung, and E. Yu, "From object-oriented to goal-

- oriented requirements analysis," *Communications of the ACM*, vol. 42, no. 1, 1999, pp. 31–37.
- [82] O. Balci, "Quality assessment, verification, and validation of modeling and simulation applications," in *Proceedings of the 36th Conference on Winter simulation*, R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, Eds. Washington, D.C., USA: Association for Computing Machinery, 2004, pp. 122–129.
- [83] D. K. Pace, "Modeling and simulation verification and validation challenges," *Johns Hopkins APL Technical Digest*, vol. 25, no. 2, 2004, pp. 163–172.
- [84] W. Feller, "The fundamental limit theorems in probability," in *Selected Papers I*. Springer International Publishing, 2015, pp. 667–699.
- [85] R. G. Sargent, "Verification, validation, and accreditation: Verification, validation, and accreditation of simulation models," in *Proceedings of the 32nd Conference on Winter Simulation*, J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, Eds. Orlando, Florida, USA: Society for Computer Simulation International, 2000, pp. 50–59.
- [86] S. Meijer, "Gaming simulations for railways: Lessons learned from modeling six games for the Dutch infrastructure management," in *Infrastructure Design, Signalling and Security in Railway*, X. Perpinya, Ed. InTech, 2012, ch. 11, pp. 275–294.
- [87] G. van Lankveld, E. Sehic, J. C. Lo, and S. A. Meijer, "Assessing gaming simulation validity for training traffic controllers," *Simulation & Gaming*, vol. 48, no. 2, 2017, pp. 219–235.
- [88] S. A. Meijer, "The power of sponges – high-tech versus low-tech gaming simulation for the Dutch railways," in *CESUN 2012: 3rd International Engineering Systems Symposium*, Delft, The Netherlands, 2012, pp. 18–20.
- [89] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 231, no. 694–706, 1933, pp. 289–337.
- [90] D. A. Middelkoop and L. Loeve, "Simulation of traffic management with FRISO," *WIT Transactions on the Built Environment*, vol. 88, 2006.
- [91] B. Rongas, S. Meijer, and A. Verbraeck, "Validity of railway microscopic simulations under the microscope: Two case studies," *International Journal of System of Systems Engineering*, p. in press.

Numerical Investigation of the Attenuation of Shock Waves by Simulating a Second, Transient Medium to Protect Vehicles Against Blasts

Burghard Hillig, Arash Ramezani, Hendrik Rothe

Chair of Measurement and Information Technology

University of the Federal Armed Forces

Hamburg, Germany

Email: burghard.hillig@hsu-hh.de, ramezani@hsu-hh.de, rothe@hsu-hh.de

Abstract—This paper describes the numerical investigations of a directed blast of 12.5 kg Pentaerithryltetranitrate and a 2.6 kg Trinitrotoluene heavy mortar grenade. The propagation of the resulting shock wave in the direction of a vehicle model is investigated. Various changes of the properties of the surrounding air have been made to produce a second, transient medium. The results are intended to answer the question as to whether and to what extent the rapid heating of ambient air offers the possibility to attenuate the amplitude of an approaching shock wave. A developed test set-up was used to simulate and record the pressure profiles by simulated pressure gauges. To get an optimal configuration, a parameter study of the properties of the second medium has been performed. As a result, as well in the case of 12.5 kg explosive ordnance as with the heavy mortar grenade, a reduction of the peak overpressure compared to the undisturbed propagation of the shock wave could be determined by about 35 % when heated rapidly to 3000 K.

Keywords—*shock wave; attenuation; blast; vehicle protection; ANSYS AUTODYN.*

I. INTRODUCTION

In this paper, the theoretical feasibility of an attenuation method for shock waves is further investigated by using numerical simulation software. The first work in this area was done by fundamental considerations, creation and simulation of a basic problem set-up [1]. Especially, the idea of forming an attenuation space out of air to protect vehicles was evaluated. This idea is based on a patent [2]. An effect of this patent was a discussion and the preceding simulation about the possibility to create a weakening mechanism against blasts like some sort of attenuation shield. Such a shield would be an effective safeguard against the threats that the armed forces are facing in today's operations.

Today's military operations differ in earlier times particularly in the fact that the stabilization of countries and peace-keeping have taken a much larger share. These operations vary in intensity, duration, environment, risk and involvement with the civilian population [3] [4]. The aim of these missions is to stabilize a failed or newly formed state. Another possibility is the reconstruction of a state after war. For that, it is necessary to establish a peaceful and stable environment for the citizen as well as for all executive elements. In order to fulfill these

missions, not only military armed forces provide security but also civilian security forces, including police, personal security and private security personnel.

Besides these stabilizing operations abroad, the number of deployments taking place within stable countries and normal peaceful conditions has increased in recent years. This development was triggered by an increase of religiously and politically motivated terrorist attacks and disruptive actions. The aim of these actions is to intimidate the population, civil aid organizations, regular army and security forces through terror. Terrorists of this kind are using explosives by means of improvised explosive devices (IEDs) to achieve their goals. This tactic has been used in many different countries and from numerous terrorist and criminal organizations [5]. Such a terrorist threat or incident can occur at any time of day with little or no warning and may result in casualties and heavy damage. Furthermore, terrorists can also use high-angle weapons with explosive ammunition, such as mortars, in order to conceal and flexibly carry out attacks on security forces.

This development results in an increased need for protection against explosive weapons for all armed and security forces. For military forces, the protection of existing armor and protective systems is already at a high level. Nevertheless, this protection is primarily designed for direct fire, splitter and shrapnel. Shrapnel are metal fragments which are flung away by the explosion of an explosive charge. But these damage effects represent only a part of the entire damage and can be handled by physical barriers such as armors.

Another damage effect is caused by the shock wave of the explosion. A shock wave is a very fast moving, strong pressure wave which is created in air by a rapid release of energy [6]. Its state variables pressure, temperature and density change almost instantaneously. The shock wave passes through surrounding air with a high velocity and with practically no resistance. If the shock wave has contact with other media during its propagation, it can be traversed by the shock wave. Physical barriers, which protect against splitter, offer little or no protection against shock waves. That is because the abrupt pressure change of the shock wave can overcome the barrier.

In order to protect people, vehicles and structures from the destructive forces of shock waves, the peak overpressure and

the rapid pressure increase of the shock wave must be reduced. Such a reduced pressure could make a major contribution to the protection of structures, vehicles and personnel of the armed forces and security forces.

In this paper, the theoretical feasibility of an attenuation method for shock waves is tested by using the numerical simulation software ANSYS AUTODYN (Canonsburg, USA). Especially the idea of forming an attenuation space to protect vehicles is evaluated.

After this introduction, Section II of this paper presents an overview about general properties of shock waves and methods of shock wave attenuation, followed by Section III about numerical simulation, spatial discretization and material models. Section IV describes the basic and advanced problem set-up for the simulation. The results of this problem set-up with respect to any attenuation will be discussed in Section V. This paper ends with a conclusion section.

A. Background and Methods

This section describes the general, theoretical properties of shock waves and is intended to provide an overview of existing ways to attenuate shock waves. In addition, the idea of an attenuation method is explained, which is examined in this paper for feasibility.

1) *General Shock Wave Properties:* When an explosion is initiated in a gas, a very rapid exothermic chemical reaction occurs. As the reaction proceeds, a part of the energy of the explosive material is released very fast. As source of energy of this reaction, a multitude of materials can be used in a wide range of aggregate states. Some well-known examples of solid explosive materials are Trinitrotoluene (TNT) and Pentaerythritoltetranitrate (PETN). But also the stored energy within a compressed, cold or hot gas can be such an energy source.

The explosive products initially expand at very high velocities and reach equilibrium with the surrounding air [7]. The properties of air as a compressible gas result in a shock wave that propagates in time. Such a wave differs from a sound wave where the displacements of the gas molecules are very small. Instead, the displacements of a shock wave have a finite value. The result is a non-linear wave in which the sound speed is related to the amplitude. The disturbances caused by the energy release travel faster than the surrounding fluid can react due to its specific sound speed. As a consequence, there is an instantaneous change of the properties of the fluid. Some examples of these properties are density, pressure and temperature. In that way, a wave front is created which consists of strongly compressed air and moves radially away from the source at supersonic velocity.

If the shock wave is caused by a detonation of an explosive charge, it is called a blast wave. This name arises from the strong wind that accompanies the wave and is felt by a stationary observer as the wave passes. In addition, the term

shock wave is generally used to describe a steep pressure gradient in which the wave pattern is not characterized in detail [8].

The behavior of a shock wave can be described by a time-dependent pressure profile. After a certain time t_a the shock wave reaches a point in space, the pressure will increase rapidly from ambient air pressure P_0 . The increase in pressure reaches a maximum peak pressure value P_{pop} and decreases after that. Figure 1 shows the described time-dependent pressure profile of an ideal shock wave as red curve. Due to the proceeding expansion of the shock wave, the wave will geometrically occupy an increasingly larger space. Because the crossed air is heated by the shock wave, a part of its energy will be dissipated as heat. That is why, the positive pressure drops exponentially. This is referred as positive phase which ends after the duration t_d . As the shock wave continues to expand, the pressure at a point in space is reduced to values below P_0 . These values asymptotically approaches the atmospheric pressure P_0 over time and enter into a state of equilibrium with the surrounding air. This is referred as negative phase.

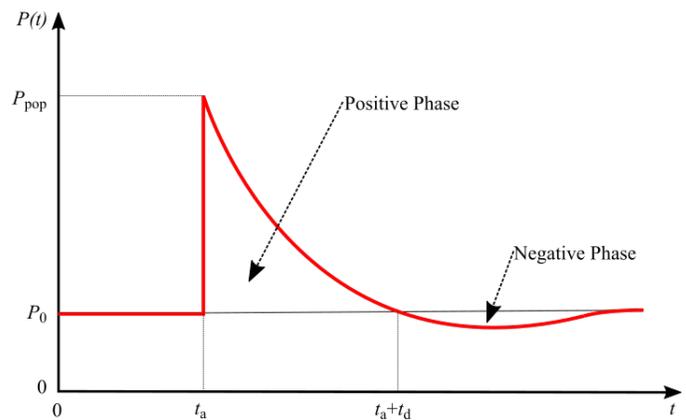


Figure 1. Characteristic time-dependent pressure profile of an ideal shock wave as a red curve. After the explosion, the wave propagates in space and reaches a point after the "time of arrival" t_a .

The easiest way to mathematically describe the pressure profile of a shock wave in air is the Friedlander equation [9]:

$$P(t) = P_0 + P_{pop} \cdot \left(1 - \frac{t}{t_d}\right) \cdot e^{-\frac{t}{t_d}} \quad (1)$$

In this equation, P_0 is the atmospheric pressure, t the time and t_d the duration of the positive phase. In this case, the positive peak overpressure P_{pop} is the value of pressure at the pressure peak subtracted by the atmospheric pressure. The Friedlander equation is valid only from the time of arrival t_a .

2) *An Overview of Shock Wave Attenuation:* The idea of shock wave attenuation to protect stationary or moving objects is not new. For structures such as bunkers or tunnel systems, shock waves can be weakened by being forced into structural channels. These channels cause multiple abrupt changes of the propagation direction of the shock wave. Also, reflections of

the wave on rough walls contribute to the attenuation by the disturbance of the unimpeded propagation [10]. If the object is endangered by shock waves on its exterior side, other methods must be used to protect the object.

Therefore, there is the possibility to attenuate the shock wave by bringing in a material between shock wave and object. This material may have the property of absorbing energy from the shock wave. An example of a realization of such a absorber is a porous material, which covers the entire surface of an object to be protected [11] [12]. A very similar protection concept can be achieved by a number of cells containing the attenuating material and placed on the surface [13]. Also, heavy but transportable attenuation shields are a feasible way to attenuate the propagation of the shock wave and thus protect persons [14].

A disadvantage of all of this methods is that the shock wave attenuating material has to be permanently applied on a structure, surface or person or it must at least be close to the object.

3) *Suggested Method of Shock Wave Attenuation:* The creation of a second, transient medium offers a dynamical method for shock wave attenuation. This second medium shall be created in a region between the site of explosion and the asset to be protected. It can be made by placing small solid particles or fine liquid droplets in a gas phase [15]. The density of air in this volume is lower but the overall density is higher. Compared to air the solid particles have a higher density and the more particles are in a volume the higher its density. A medium created in this way has a higher density than the surrounding air. It can be observed that a crossing shock wave is attenuated through this dust-gas-suspension [16].

Such an attenuation can also be achieved if the second medium is not a suspension. This can be accomplished through rapid heating of a volume of air between the shock wave and the object [2]. By that, the heated air creates a medium, which differs in density, temperature or composition from the surrounding air. Such a medium should reduce the energy density of the crossing shock wave by reflecting a part of the shock wave on the interface between the media. Another part of the shock wave could be scattered away from its original propagation direction. In addition, the wave front can be influenced to diverge within the second medium and gets partly absorbed. In this way, the entire shock wave should be weakened.

To accomplish the attenuation of an approaching blast wave in reality, it is necessary to have a system which offers the capability to detect an explosive event, determine its direction and propagation velocity. Subsequently, the system has to create a second, transient medium to protect an asset. This second medium has to be placed between shock wave and the object. Furthermore, the medium should be formed by rapidly heating a zone of the surrounding, ambient air.

From a technical point of view, such a system should include a sensor for detecting the source of the shock wave, a computer

system for processing the detected signal, and a triggering countermeasure communicating with the detection system. In Figure 2, a sketch of this system is shown based on the previously cited patent. The asset to be protected can be a military or a civil safety vehicle. The described countermeasure has the task to create a second medium by heating the surround air, which depends on the position relative to the asset to be protected, the spatial expansion and the composition. The heating of this intermediate region can be realized through an electromagnetic arc. Through this electrical arc, an electric current can flow through air by using a conductive path which is made by the dielectric collapse of the ionized air. Other possible ways to archive this heating are by focused microwaves or laser beams.

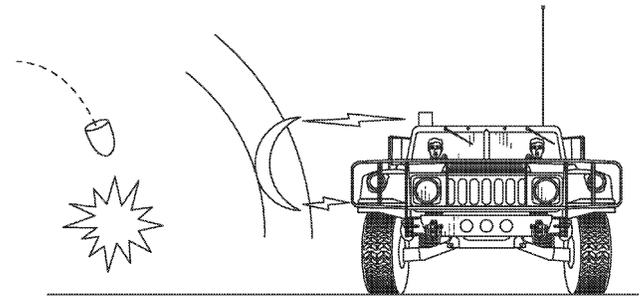


Figure 2. Sketch of a vehicle and an explosive event. The vehicle protects itself by activating a countermeasure which creates a second, transient medium by heating the ambient air between the explosion and the vehicle. Adapted from [2].

The surround air can also be heated without an electric arc by using multiple conductive strips or threads. Starting from an energy source, a very strong electric current is sent into the conductive material, so that it evaporates and strongly heats the air.

In principle, the presented method for attenuating shock waves can be applied to all types of vehicles, fixed structures, airplanes and also ships. However, for the subsequent investigations, the technical realization of the attenuation space is not a priority. Instead, numerical simulations with the software ANSYS AUTODYN are used to investigate the theoretical feasibility and verifiability of a blast wave attenuation by using the described effect of a rapidly, heated region of air.

II. NUMERICAL SIMULATION

This section explains the basic principles of numerical simulation and describes the material models used in this work.

A. Spatial Discretization

When it comes to solving problems related to the release of huge amounts of energy in very short time then different

solutions are possible. A good possibility to describe such processes offer analytical solutions. Unfortunately, their applicability is limited to problems with simple geometries and few boundary and initial conditions. Numerical simulations, on the other hand, offer a solution to common problems and deal with difficulties related to geometry [17].

The basic physical model of a numerical simulation can be attributed to physical conservation laws, the equation of state and the constitutive model. Partial differential equations for energy, momentum and mass constitute the physical conservation laws. Furthermore, the equation of state connects the internal energy respectively temperature and density of a material with its pressure. As a result, changes in the density and irreversible thermodynamic processes such as shock-like heating can be considered. In addition, the constitutive model contains the effects of the material to be simulated and describes the effect of change in shape and material strength properties.

Together, these equations build a set of coupled, time- and space-dependent, highly non-linear partial differential equations. These governing partial differential equations must be solved in both time and space domain. The solution can be achieved via computer simulations with an explicit method, in which the solution is expressed at a given time as a function of the system variables and parameters, without stiffness and mass matrix requirements. The computation time for every time step is low, but may require numerous time steps for a complete solution. The solution for the space domain can be achieved by using different spatial discretization methods, like Lagrange [18], Euler [19], Arbitrary Lagrange-Euler [20] or a mesh free method [21]. Each of this methods has its unique advantages but also limitations. So, normally there is no single technique, which handles every part of the problem correctly [22].

In the presented work, two of these methods are used for spatial discretization: the Lagrange and Euler method. The Lagrange method is used to model objects. In return, the multi-material Euler solver is used to model air and the explosive.

1) *Lagrange*: The Lagrange method divides an object into a spatial grid, which is bound rigidly to the object and moves with it. The material component within an element does not change. If forces are acting on a node of an element, the node is displaced and passes the forces to its neighboring nodes. The behavior of this configuration can be imagined similar to a spring-mass system. The described effect causes a deformation of the mesh. Boundary and initial conditions can be applied easily, so that the nodes of the boundary elements of an object stay unchanged. Also, very clear material boundaries exist in the Lagrange method. For the space outside the object, there is no extra mesh needed and mass conservation is automatically satisfied.

An example of this process is shown in Figure 3. A target and a bullet like object are modeled with a blue and a green material. The bullet (blue) has an initial velocity in the direction of the target. The left side shows the situation before

the impact of the bullet. During the impact process the target (green) gets deformed and the deformation of the target is calculated as a result of the deformation of the Lagrange mesh of both objects at each time step.

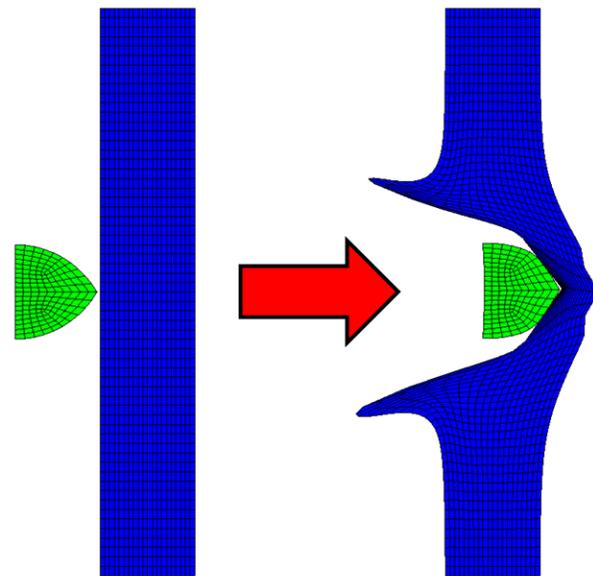


Figure 3. Lagrange method example: A bullet (green) and a target (blue), both with a Lagrange mesh. The left side shows the situation before and the right side after the impact of the bullet on the target. Adapted from [23].

In general, the Lagrange method is best suited for complex geometries like structures, projectiles and other solid bodies. A disadvantage of the Lagrange method is that strong mesh deformations may occur at high loads. Such a distorted mesh element can negatively influence the time-dependent solution because the time step is proportional to the size of the smallest element within the Lagrange mesh.

2) *Euler*: The Euler method differs in this way that the coordinates of the nodes are fixed and they fill the entire space. The material flows time-dependent through the mesh and changes the properties of the elements, while the spatial coordinates and nodes stay unchanged. That is the reason why no mesh distortion can occur at the Euler method. In contrast to the Lagrange method, the boundary nodes do not necessarily coincide with the material boundary conditions. This can result in difficulties in the application of boundary and interface conditions.

Figure 4 shows a bullet and a target modeled in the Euler frame. In contrast to the Lagrange frame, the whole simulation space consists out of a mesh. Part of this is the white space on both sides of Figure 4. Other parts of the simulation space are filled with material matching the objects they are meant to represent. After the impact, the material is spread over the mesh. The result shows a situation which fits Figure 3.

In general, the Euler method is used to represent fluids and

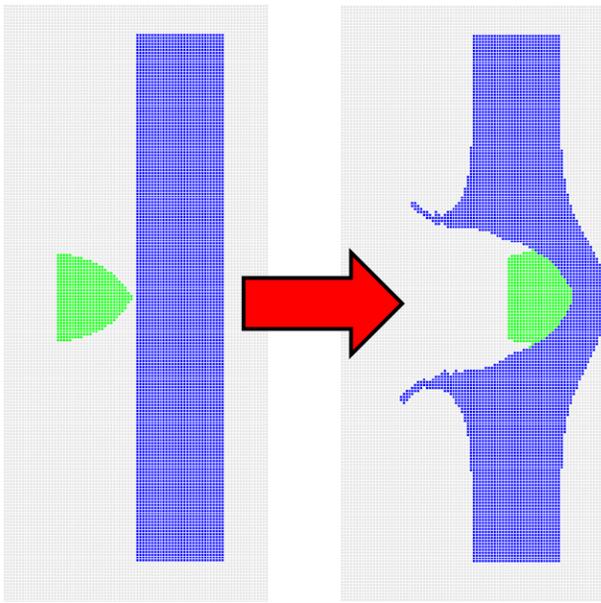


Figure 4. Euler method example: A bullet (green), a target (blue) and a (white) simulation space are modeled with an Euler mesh. The left side shows the situation before and the right side after the impact of the bullet on the target. Adapted from [23].

gases, especially for modeling the propagation of a blast wave as result of an explosion. If a pure simulation of solids is needed, the Euler method has the disadvantageous effect, that the stress tensor and the history of the material has to be transported through the mesh. In this case, Euler needs more computation performance and smaller elements to model the appearing blast wave correctly.

B. Material Models

1) *Explosive*: PETN and TNT serve as explosive ordnance through which the blast wave is generated at the beginning of the simulations. As equation of state, the Jones-Wilkins-Lee (JWL) equation is used. It describes the blast wave pressure as a function of the relative volume $\eta = \rho_0/\rho$ and the internal energy E [24]:

$$P = A \left(1 - \frac{\omega \cdot \eta}{R_1}\right) \cdot e^{\frac{R_1}{\eta}} + B \left(1 - \frac{\omega \cdot \eta}{R_2}\right) \cdot e^{\frac{R_2}{\eta}} + \omega \cdot \rho \cdot E \quad (2)$$

In this equation, A , B , R_1 , R_2 and ω are material constants, which were determined in dynamic experiments.

Four different material models are defined for PETN in AUTODYN. These four models differ only in their densities $\rho_{0.88} = 0.88 \text{ g/cm}^3$, $\rho_{1.26} = 1.26 \text{ g/cm}^3$, $\rho_{1.50} = 1.50 \text{ g/cm}^3$ and $\rho_{1.77} = 1.77 \text{ g/cm}^3$. Simulations of explosives with these densities lead to nearly identical pressure profiles [25]. That is why the PETN material model of $\rho_{1.77}$ has been used in the first part of this work.

In the second part, TNT is used for simulation. Its behavior can be modeled with the same equation of state as for PETN. The JWL-equation will then have TNT specific material constants and the density of TNT $\rho_{\text{TNT}} = 1.63 \text{ g/cm}^3$ is used for correct calculations [24].

2) *Ideal Gas*: Air is a gas mixture mainly consisting of diatomic nitrogen and oxygen. In addition, a multitude of mono- or poly-atomic trace gases are contained in air. Furthermore, the ideal gas equation is a suitable equation of state for modeling air in numerical simulations:

$$P = \rho \cdot R_s \cdot T \quad (3)$$

In this equation, P is the pressure, ρ the density of the gas, R_s a specific gas constant and T the temperature. An alternative but equivalent formulation of the previous equation is used in ANSYS AUTODYN for its internal calculations [24]:

$$P = (\gamma - 1) \cdot \frac{\rho}{\rho_0} \cdot E \quad (4)$$

In this alternative equation, γ is the ratio of the specific heat capacities at constant pressure $c_p(T)$ and volume $c_v(T)$. This ratio has the value $\gamma = 1.4$ under the conditions of ambient temperature at $T_{\text{ref}} = 288 \text{ K}$ and atmospheric pressure of $P_{\text{atm}} = 101.3 \text{ kPa}$. Based on these values, the reference density is $\rho_0 = 1.225 \text{ kg/m}^3$ and the internal energy has the value $E = 253.4 \text{ kJ/m}^3$.

A complete set of parameter-based data for a large interval of temperature values can be found in measurement-based tables [26] [27]. The corresponding values for γ can be extracted from this data. With these values, the internal energy E of a gas at a certain temperature T can be calculated by following equation:

$$E = \int_{T_{\text{ref}}}^T c_v(T) \cdot dT \quad (5)$$

Table I is made up of the empirical and calculated data for a temperature range of 288 K to 3000 K. The density was calculated as a function of the temperature according to (3) at constant atmospheric pressure P_{atm} .

Since air is a gaseous material, it does not suffer shear stresses or negative pressures. That is the reason why no strength or failure model is necessary for the numerical simulation of air.

It should be noted that a heating process of air based on (3) is realized under the conditions, that either volume or pressure stay constant. A quick heating with constant volume will result in a very strong change of pressure. On the other hand, an expansion of the volume, which equals a decrease in density, leads to rapid heating with constant, ambient pressure.

3) *Structural Material*: Structured parts will be simulated by using the Lagrange method and the material steel-1006. The further material properties of steel are irrelevant, because

TABLE I. VALUES FOR DENSITY, INTERNAL ENERGY, HEAT CAPACITY AT CONSTANT VOLUME AND PRESSURE OF AIR AT VARIOUS TEMPERATURES.

Temperature [K]	Density [g/cm ³]	Internal Energy [J/kg]	C_V [kJ/kg K]	C_P [kJ/kg K]
288	0.001225	2.068E+05	0.7173	1.0042
300	0.001177	2.143E+05	0.7177	1.0050
400	0.000882	2.865E+05	0.7262	1.0130
500	0.000706	3.600E+05	0.7423	1.0290
700	0.000504	5.127E+05	0.7876	1.0750
1000	0.000353	7.597E+05	0.8538	1.1410
1200	0.000294	9.341E+05	0.8873	1.1740
1500	0.000235	1.206E+06	0.9239	1.2110
1700	0.000208	1.393E+06	0.9419	1.2290
2000	0.000176	1.679E+06	0.9630	1.2500
2200	0.000160	1.872E+06	0.9731	1.2600
2500	0.000141	2.165E+06	0.9866	1.2740
2700	0.000131	2.364E+06	0.9942	1.2810
3000	0.000118	2.663E+06	1.0040	1.2910

it is only used to fill the object as a rigid body. An interaction between blast wave and object does not take place.

III. PROBLEM SET-UP

The following subsections describe the created problem set-up for the numerical simulation. Each element of the simulation space and all parts of the blast process are explained. The goal of this section is to get an understanding of how reality can be modeled in numerical simulations and what simplifications have been done to achieve that.

A. Basic Set-Up

1) *Undisturbed Propagation:* In order to measure the effects of an attenuation space for blast waves, a simulation set-up was developed, which ensures uniform and comparable conditions between individual simulations. This set-up is designed in such a way that the undisturbed propagation of the blast wave and its positive peak overpressure P_{pop} can be measured at various points in the simulated space. Subsequently, the individual parameters of the simulation are to be varied in order to be able to compare the effects on the pressure profile with the undisturbed case. Parameters to be changed are the thickness of the attenuation space, its temperature, density, pressure, position and shape. Other components of the simulation, such as the type of explosive and the type of space to be viewed, remain unchanged.

The approximate spatial configuration of the simulation is based on a certification standard for non-destructive requirements on special-safety vehicles. In order to validate the attenuation space, it is important to check the change of pressure on the vehicle by gauges. In this first view, the direct impact of the blast wave on the vehicle is negligible. That is the reason why the vehicle can be simulated as a rigid body like a dummy. This dummy vehicle has the shape of a sport

utility vehicle (SUV), which is a good example for a security vehicle and is similar to military off-road vehicles.

The space, in which the blast wave should propagate, will be modeled with the Euler method. The Euler Multi-Material solver is used to fill a space with air under ambient temperature and atmospheric pressure. The created space is designed in such a way that it can represent all important elements of the propagation of the blast wave in the direction of the vehicle. It has a length of $l = 13.00$ m, a width of $b = 6.00$ m and a height of $h = 3.00$ m. The individual hexahedral elements of the mesh in the Euler space have an edge length of 0.05 m. Accordingly, the space is filled with 1.872 million elements. The entire set-up of the simulation is shown in Figure 5.

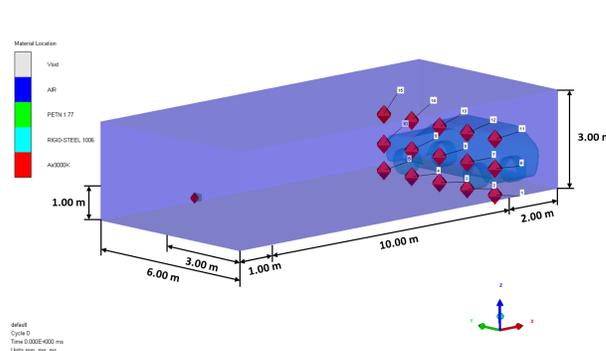


Figure 5. The simulation space filled with different materials and a vehicle model from an oblique, lateral perspective with gauges and PETN explosive.

The entire simulation space is mainly filled with air (blue space) at ambient condition and with a vehicle model (dark blue). The view on this simulation space is from an oblique, lateral perspective. The sides of the simulated space are provided with boundary conditions, which allow the blast wave to flow out. In this way, a disturbing reflection of the blast wave on these sides is avoided. Only the floor of the space is excluded from this condition, so that the blast wave is reflected. By applying this boundary condition, the computation effort is reduced and the same result is achieved as if the room was an infinite space.

At the end of the long side of this space, the vehicle is positioned in such a direction that it will be hit on the side by the blast wave. In Figure 5, the PETN explosive with its detonation point is located to the left of the front of the space. The explosive is placed in a distance of 10 m away from the vehicle, exactly centered and hovering in 1 m height. The hovering state is realized in the simulation just by placing the explosive at this coordinate. Because of typically very short durations in high-speed dynamics every influence of gravitational forces on the position of the explosive can be neglected. In reality, the localization of the explosive would be realized by a platform of easily destructible material, which holds the explosive. The volume of the explosive is chosen in such a way that it equals an amount of 15 kg TNT. The TNT

equivalent of PETN is approximately 1.2, that is why 12.5 kg of PETN are required [28]. In accordance with the certification standard, the explosive is chosen in the shape of a cube. The edge length a of this cube for a given mass $m_{\text{PETN}} = 12.5$ kg and density $\rho_{1.77}$ can be calculated by its volume V :

$$a = \sqrt[3]{V} = \sqrt[3]{\frac{m_{\text{PETN}}}{\rho_{1.77}}} \approx 0.19 \text{ m} \quad (6)$$

The detonation point for the explosive is centered on the surface which is located on the side facing away from the vehicle. This results in a blast wave moving in the direction of the vehicle. Left side to the front of Figure 5 shows the PETN cube with its red detonation point.

The measurement of the pressure profile of the blast wave is done by fifteen pressure gauges. In Figure 5, the red dots on the right side represent these gauges at which the pressure can be measured. They are located above each other, three rows of five gauges directly in front of the rigid vehicle. Each gauge can collect data of various physical properties during the entire time of simulation.

The presented set-up can be used to perform simulations of the undisturbed propagation of blast waves.

2) *Disturbed Propagation:* In a next step, an attenuation space is placed directly in front of the vehicle. This space consists of front and back faces with constant area and a variable thickness in the direction of the blast. Furthermore, position, shape, temperature, density and pressure are changeable. This set-up with the newly introduced attenuation space is shown in Figure 6.

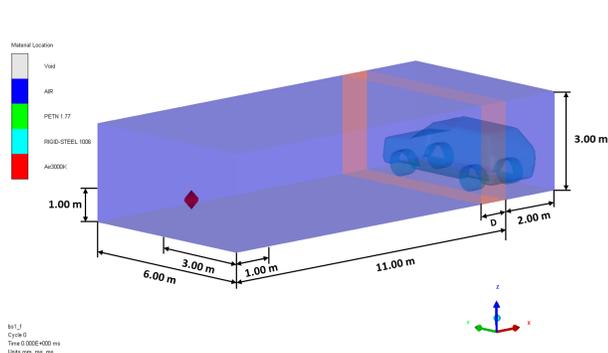


Figure 6. The simulation space filled with different materials and a vehicle model from an oblique, lateral perspective with a detonation point on the left side. The reddish space on the right side represents the attenuation space.

The entire simulation space is filled with different materials and a vehicle model from an oblique, lateral perspective. The mainly blue space in this figure is filled with air at ambient condition. The reddish space on the right side represents the attenuation space. This space has a thickness D , which can be set arbitrarily in the direction of the blast. Furthermore, shape and density are variable. The previously described gauges exist

in this set-up as well and they record changes in pressure during the simulation. For the sake of clarity, the gauges are hidden in this figure.

B. Advanced Set-Up

The simulation scenario described before is a very basic and artificial attempt to estimate the effect of a blast attenuation mechanism. However, due to the high strength of the 12.5 kg PETN explosive used in the previous set-up, the set-up shows an upper limit of its attenuation capability.

To validate a more realistic scenario the basic set-up is modified. A huge PETN explosive ordnance is a good choice for the certification of the demolition of a vehicle but in reality other threats like high-angle weapons with explosive ammunition are more plausible. For example, an incoming mortar grenade next to a security vehicle. Such a mortar grenade is a weapon not only used by military units but also by insurgents and terrorist. It is a very simple weapon with cheap ammunition, which are distributed vastly due to the cold war. It also satisfies the requirements for a flexible weapon as described in the introduction. The actual form of this weapon was developed to support infantry units as a light and mobile weapon [29]. It fills the gap between artillery and grenades, so that the troops can adjust flexibly and quickly to possible threats.

The advanced scenario will use a representation of a mortar grenade as explosive instead of the huge 12.5 kg PETN explosive ordnance. The remaining configuration stays the same.

1) *Mortar Ammunition:* Mortars are available as lighter or heavier versions. The lighter version normally has a caliber of 81 mm with a range of about 5000–6000 m. A heavy mortar has a increased weight but it can fire its 120 mm shells to a distance of about 7000–8000 m. High firing distances and light overall weight of about 50–120 kg make the mortar a flexible and easy transportable weapon [29].

As shown in Figure 7, the body of a mortar grenade consists of a fin stabilized tail and a shell. Essentially, the shell of the mortar grenade carries the explosive with the fuze at the top of the shell. Depending of what kind of fuze is used it is possible to let the grenade explode by impact or above the ground [30]. The grenade shown in Figure 7 is a light version of a Russian 120 mm grenade, which should be used as an example for the following simulation. The heavy grenade has an explosive filler with weight of $m_{\text{grenade}} = 2.6$ kg TNT and a overall length of $l_{\text{overall}} = 0.665$ m [31].

2) *Modeling a Mortar Grenade:* To model such a grenade in ANSYS AUTODYN, the volume of the explosive filler is needed. Its value can be calculated by using the simple relationship between density ρ_{TNT} , volume V and mass m :

$$V = \frac{m_{\text{grenade}}}{\rho_{\text{TNT}}} \approx 1.58 \cdot 10^{-5} \text{ m}^3 \quad (7)$$

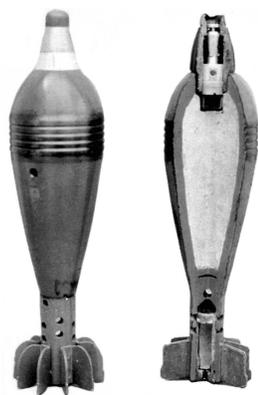


Figure 7. Illustration of a light 82 mm high explosive mortar grenade as used during cold war and today. The right side shows a cross-section with the high explosive material filler in the middle and the fuze on top of it [31].

For the sake of simplicity, in the simulation of the mortar grenade this volume can be represented by a cuboid with square base. The edge length of the square base is $a = 0.10$ m and $h = 0.15$ m for its height or longest dimension. These dimensions lead to a volume which satisfies the calculation from equation (7).

For the simulation, an outer shell of the mortar grenade is not needed because the release of shrapnel is not relevant. Only the blast wave is needed to get an impression about the interaction with the attenuation space. The fuze of the grenade can be modeled as a detonation point on top of the cuboid. The described design is realized in Figure 8.

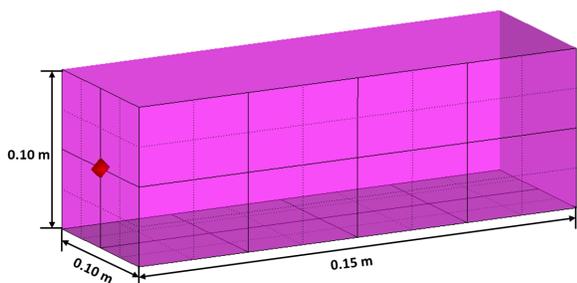


Figure 8. Simplified computer model of a mortar grenade. The space in the magenta color is filled with TNT and the red point on top is the detonation point.

The modeled grenade in Figure 8 is filled with TNT, which is represented by the color magenta. It also has a detonation point on top, shown as red dot. This simple representation of the grenade shown in Figure 7 meets all requirements to allow a simulation.

3) *Simulation Space*: The modeled mortar grenade will be placed in the simulation space instead of the huge PETN

explosive ordnance. It will be placed in such a way that the normal component of the surface, where the detonation point is located, points toward the ground of the simulation space. This imitates the behavior of a real mortar grenade because its velocity direction would point toward the ground before impacting and exploding.

In Figure 9, the mortar grenade model from Figure 8 is part of the advanced simulation set-up. Again, the grenade is represented by a magenta color with the detonation point pointing to the ground.

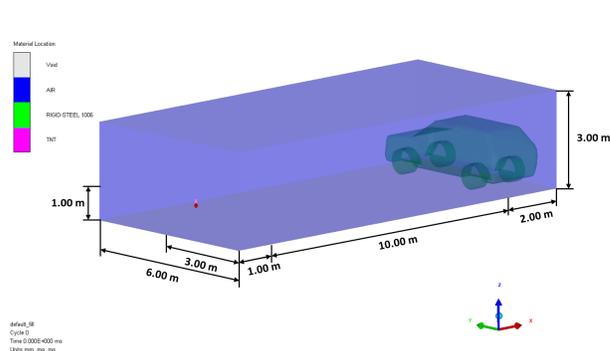


Figure 9. The simulation space filled with air and a vehicle model from an oblique, lateral perspective. On the left side of the space a simplified mortar grenade model with a detonation point (red dot) was inserted.

The configuration of this set-up is the same as in the basic case. Only the explosive ordnance is replaced. With this set-up the undisturbed propagation of the blast wave and a propagation through an attenuation space can be investigated. This offers the possibility to check if the attenuation of the blast wave will be achieved in the case of a smaller and more realistic explosive threat for the vehicle.

IV. RESULTS AND DISCUSSION

The simulation results between the undisturbed and the attenuated blast wave propagation with the basic and advanced set-up are described below. The results are compared and discussed.

A. Undisturbed Propagation of the Basic Set-Up

The numerical calculations of the undisturbed blast wave propagation agree well with the expected ideal behavior. After the start of the simulation, the detonation of the PETN explosive leads to a successive conversion of the explosive. This results in a blast wave, which propagates radially away from its source. A part of the expanding blast wave is reflected on the ground and overlaps with the wave front. Hence, spaces of the blast wave with higher pressure are created, which can be observed near the ground.

Figure 10 shows this situation in the simulated test set-up. The blast wave is represented by vector arrows of the velocity

component in a color coding. The color changes from red to blue, as does the speed from high to low. The shown blast wave expands radially away from the explosive ordnance towards the vehicle. A part of the wave front is superimposed by the reflection on the ground.

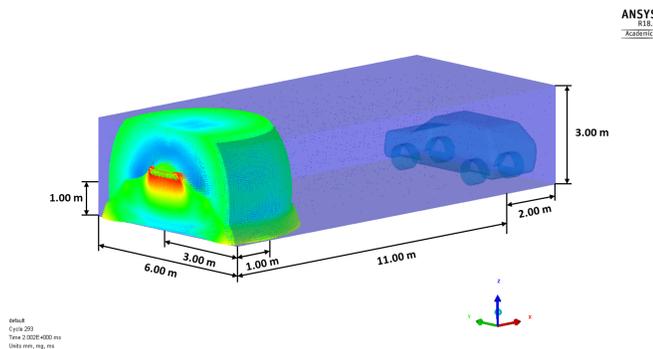


Figure 10. The simulation space filled with different materials and a vehicle model from an oblique, lateral perspective. A blast wave expands radially outwards from the left front side towards the vehicle.

If the blast wave reaches the vehicle, the pressure gauges will record the pressure change. For the evaluation, the data of gauges 3 (g3), 8 (g8) and 13 (g13) are used. These gauges are located exactly in the middle of the dimension along the y-axis of the simulation set-up. Along z-direction they are located at the ground, in middle height and on top of the vehicle. The pressure profile of the recorded data is shown in Figure 11. The overall profile of the measured pressure is in good accordance to the profile of the ideal shock wave from Figure 1. The peak overpressure values can be extracted from the recorded data. It is important to notice that these peak overpressure values are valid with respect to the atmospheric pressure of 100 kPa.

The gauges measured a peak overpressure of $P_{g3} = 62.87$ kPa at the bottom, $P_{g8} = 62.45$ kPa in the middle and $P_{g13} = 62.94$ kPa on top. Vehicles that are hit by a blast wave with a pressure amplitude of about 62 kPa are overturned and heavily damaged. Buildings cannot resist this overpressure as well: wood framed buildings collapse, steel framed buildings receive serious damage and reinforced concrete structures suffer severe damage [32] [33].

B. Heating under Atmospheric Pressure of the Basic Set-Up

For the next step, a part of the simulation space was filled with a cuboid-shaped volume of modified air at which the feasibility of the blast wave attenuation can be verified. The volume of this space is defined as the entire height and width of the simulation space. The thickness of the space was set to $D = 1.00$ m and placed in a distance of d away from the vehicle, which in this case equals the amount of the thickness. Through that, the modified volume fills the free space without touching any part the vehicle. To emulate a rapid heating

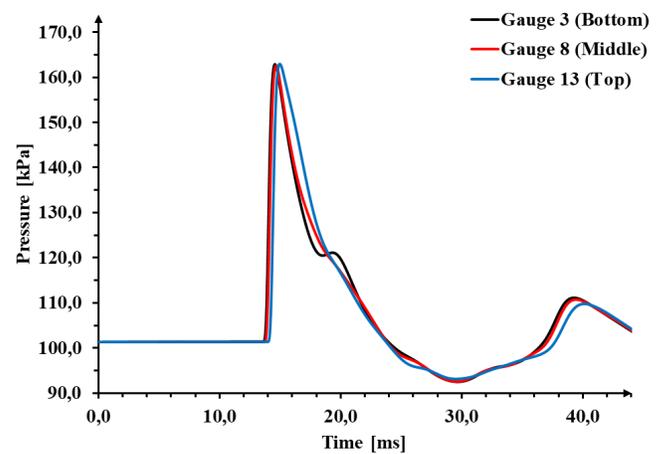


Figure 11. Simulated time-dependent pressure profile of the blast wave of the simulation set-up. After its release through the explosion, the wave propagates spatially and reaches the vehicle after $t_a \approx 14$ ms. In the ideal profile, a positive and negative phase of the pressure is visible.

under atmospheric pressure, the ambient air was replaced by air which matches the properties and internal energy of air at a temperature of $T = 3000$ K according to Table I.

Three recorded pressure profiles of the blast wave, which has to cross such a volume in the simulation space, are shown in Figure 12.

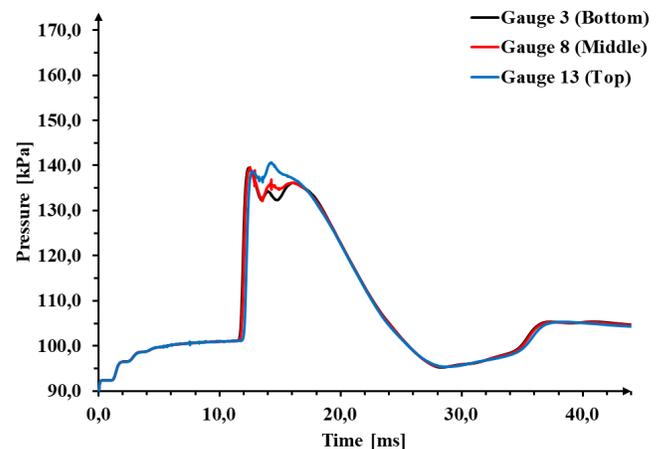


Figure 12. Simulated time-dependent pressure profile of the blast wave after crossing the attenuation space. The pressure profile shows an attenuation in the direction of the maximum pressure.

At the start of the pressure recording, the used method of heating will result in a slight vacuum because parts of the adjacent ambient air flow into the heated region because of the reduced density. The pressure stabilizes until the time of arrival of $t_a \approx 14$ ms. When the blast wave reaches the gauges, a sudden increase in pressure occurs like in the undisturbed

case. In contrast, the measured peak overpressure at the gauges is lower and has the following values $P_{g3,heat} = 40.87$ kPa, $P_{g8,heat} = 41.98$ kPa and $P_{g13,heat} = 45.79$ kPa.

In direct comparison, the relative differences of these peak overpressure values are $\Delta P_{g3,rel} = 35.0\%$, $\Delta P_{g8,rel} = 32.8\%$ and $\Delta P_{g13,rel} = 27.2\%$ lower than the undisturbed case. These lower peak overpressure values are the consequence of the propagation of the blast wave through the attenuation space. Because of that, the object to be protected experience lesser damage and an attenuation of the blast wave through a second, transient medium is provable in the simulation. Furthermore, it can be seen in the pressure profile of Figure 12 that the pressure stays at a lowered level where the sharp peak used to be. Finally, the pressure drops off similar to the undisturbed case.

The three different gauges show small differences in their individual profiles. In contrast to the undisturbed case, the peak pressure rises from bottom to top gauge. This can be an indication for a disturbance of the radial, symmetric propagation of the blast wave. An explanation can be the diffraction of a part of the blast wave in the direction of the top gauge. Another reason can be that a fraction of the propagating blast wave is reflected from the ground in an upper direction towards the top gauge.

Because the attenuation space provided first results, the influence of its parameters temperature, thickness and distance was verified by a parameter study. The first changed parameter was the position d of the attenuation space relative to the vehicle. The change was applied in steps of $\Delta d = 1$ m. To compare the difference, the peak overpressure values of the undisturbed case were related with the attenuated values as normalized overpressure values. Figure 13 shows these normalized overpressure profiles.

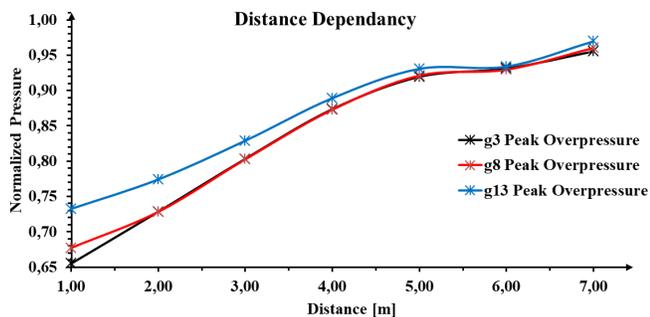


Figure 13. Distance-dependent normalized overpressure profile of the simulated attenuation space. The simulation data of the bottom, middle and top gauges are plotted and connected with smoothed curves.

The normalized overpressure rises with increasing distance from the vehicle. The individual records of the three gauges are similar to each other. But at near distances $d < 4.00$ m, the profile of the top gauge (g13) shows a recognizable difference to the other two gauges. This difference decreases

with increasing distance d . This behavior can also be an indication for diffraction of the blast wave or the reflection of parts of the wave from the ground towards the top. In conclusion, to get a high attenuation of the blast wave, the attenuation space should be placed as near as possible to the asset to be protected.

The next parameter variation was done under the same conditions as before. With the attenuation space near the vehicle, the thickness D of the space was altered. The resulting normalized overpressure profile is plotted in Figure 14. It shows that the normalized overpressure decreases with increasing thickness and stabilizes at values of $D > 1.50$ m. For best results, the thickness should be at least $D = 1.50$ m.

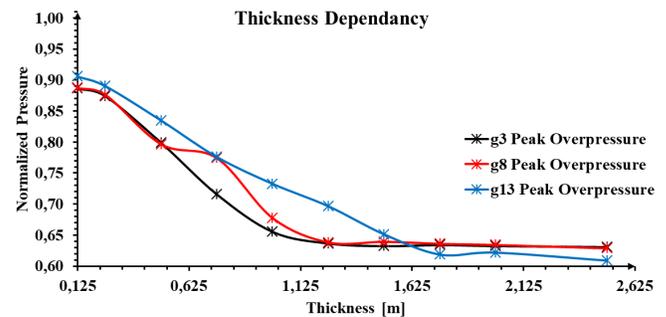


Figure 14. Thickness-dependent normalized overpressure profile of the simulated attenuation space. The simulation data of the bottom, middle and top gauges are plotted and connected with smoothed curves.

In a next step, the temperature was changed as a parameter. To do that, the same conditions as before apply with a change in temperature. Figure 15 shows these profiles.

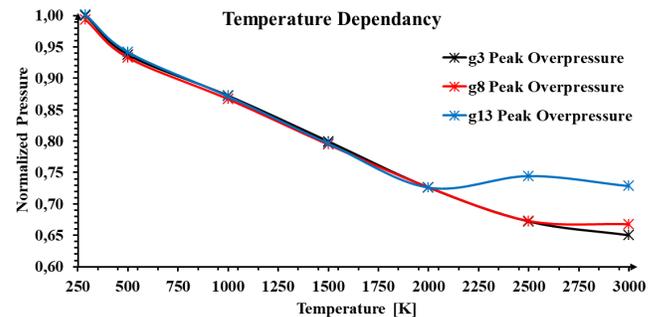


Figure 15. Temperature-dependent normalized overpressure profile of the simulated attenuation space. The simulation data of the bottom, middle and top gauges are plotted and connected with smoothed curves.

In the case of changed temperature, the normalized overpressure decreases with increasing temperature. The measurements of the three gauges behave alike to temperatures of $T > 2000$ K. From that temperature on, the values of middle and top gauge differ from the bottom gauge. Again, this can be indicating a diffraction of the blast wave from bottom to top.

Overall, the attenuation is the better the higher the temperature is.

The evaluation of this parameter study shows the following values as the optimal setting for the attenuation space in the simulation set-up: Thickness $D = 1.50$ m, distance $d = 1.50$ m, temperature $T = 3000$ K and the properties of air at this temperature according to Table I. With this setting of parameters and the results, the best overpressure values in the simulation could be extracted: $P_{g3,best} = 39.43$ kPa, $P_{g8,best} = 39.61$ kPa and $P_{g13,best} = 40.67$ kPa. The relative differences to the undisturbed case correspond to: $\Delta P_{g3,best} = 37.3\%$, $\Delta P_{g8,best} = 36.5\%$ and $\Delta P_{g13,best} = 35.4\%$.

In contrast to the undisturbed case, a lesser pressure difference of about 40 kPa results in less drastic consequences for the vehicle being hit. In this magnitude, there is no overturning of vehicles and a complete destruction of the vehicle is no longer to be expected. It is shown that under the correct application of the attenuation space in the simulation a noticeable attenuation of the blast wave can be detected.

Finally, it should be noted that the simulation differs from the real conditions in which a volume can be heated. At first, the heated and simulated space is discretely separated from the environment. Any change of the properties in the attenuation space does not change anything in the surrounding ambient air before the simulation starts. Second, the heating takes place in the case of constant atmospheric pressure in a very short time. In reality, interactions with the surrounding air, in particular diffusion, are to be expected during such a heating process. Additionally, the increase in temperature of $T > 3000$ K leads to an ionization of air. Such an ionization is accompanied by a change in density, which cannot be shown in the simulation. It is also important that steel can melt at such high temperatures. Therefore, the attenuation space must not be in direct contact with the object to be protected. In reality, it is necessary to examine how such rapid heating of ambient air affects the asset, as well as people and objects nearby.

C. Results of the Advanced Set-Up

The simulation of the modified set-up starts with the ignition of the modeled mortar grenade. This leads to a similar blast wave behavior as in Figure 10. The shown blast wave expands radially outwards from the left front side towards the vehicle. Again, a part of the wave front is superimposed by the reflection on the ground. When the blast wave arrives at the vehicle, changes of pressure will be recorded by pressure gauges. As before, only the gauges located in the middle of the ground, in middle height and on top of the vehicle are used for collecting and evaluating data.

The pressure profile of the collected data in the undisturbed case caused by the mortar grenade is shown in Figure 16.

Compared to the ideal shock wave behavior from Figure 1 and the undisturbed propagation from Figure 13 there are some differences. The rise of the amplitude is not as steep as in the

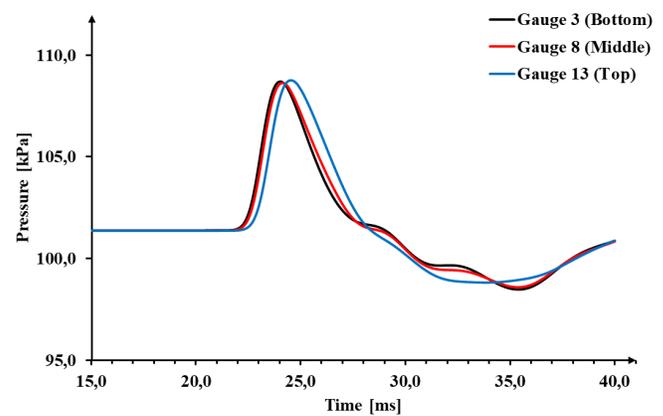


Figure 16. Simulated time-dependent pressure profile of the blast wave of the simulation set-up. After the explosion of the mortar grenade, the wave propagates spatially and reaches the vehicle after $t_a \approx 22$ ms. As in the ideal profile, a positive and negative phase of the pressure is visible.

both other cases. The reason for this is that the blast wave has already slowed down and is increasingly turning into a compression wave. The cause of the slowdown is the smaller amount of explosive at the same distance as before. Also, the amplitude of the blast wave is not as high as with the huge PETN explosive.

The results are peak overpressure values of $P_{g3,mor} = 8.70$ kPa, $P_{g8,mor} = 8.63$ kPa and $P_{g13,mor} = 8.76$ kPa. At these overpressure values the following damages are expected: Glass failure, houses are made inhabitable, slight distortion of the steel frames of clad buildings and slight to serious injuries of persons [32] [33]. The amplitude is not strong enough to overturn a vehicle. It is to be expected that a security vehicle will withstand such a blast wave.

To get a comparable result about a possible attenuation of the blast wave, the set-up was provided with an attenuation space. Like before a cuboid-shaped volume of modified air with the same properties ($T = 3000$ K) was used. The recording of the resulting pressure profile is shown in Figure 17.

From slight underpressure at the beginning, the pressure stabilizes in Figure 17 until the time of arrival of $t_a \approx 22$ ms. When the blast wave reaches the gauges, a steep but not too sudden increase in pressure occurs. This indicates that the wave has been weakened. There are also some minor peaks in the profile of gauge 8 and 13. An explanation could be an turbulence in the propagation of the blast wave. However, because of the regular, common intervals at the profiles of both gauges it might be an error in the calculations of the simulation software. Compared to each other only the profile of gauge 13 has slightly higher values.

In contrast to the undisturbed case shown in Figure 16 the measured peak overpressure at the gauges is lower and has the following values $P_{g3,mor,h} = 5.75$ kPa, $P_{g8,mor,h} = 5.77$ kPa and $P_{g13,mor,h} = 6.12$ kPa. In direct comparison with the

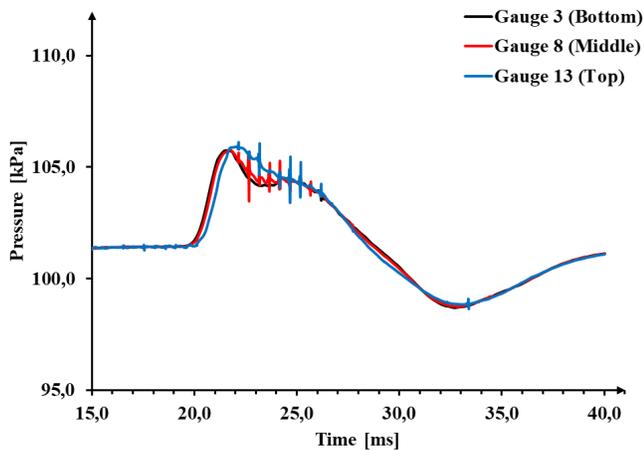


Figure 17. Simulated time-dependent pressure profile of the blast wave after crossing the attenuation space. The pressure profile shows an attenuation in the direction of the maximum pressure.

values $P_{g3,mor}$, $P_{g8,mor}$ and $P_{g13,mor}$ the relative differences of these peak overpressure values are $\Delta P_{g3,mor,rel} = 33.9\%$, $\Delta P_{g8,mor,rel} = 33.2\%$ and $\Delta P_{g13,mor,rel} = 30.1\%$ lower. Like in the set-up before an effect of the attenuation space is measurable.

V. CONCLUSION

Assets of armed and security forces are threatened by terrorists and insurgents, especially by explosive weapons. The armor of these assets effectively protects them against splinters and shrapnel, but not against the blast wave of the explosion. Therefore, in order to reduce the danger of explosions, the blast wave has to be weakened before reaching the asset. This can be achieved by rapidly heating a region of normal ambient air to produce a second, transient medium at which the blast wave is attenuated.

In this work, a simulation set-up was developed for this purpose, which could be used to check whether the heating of a volume of air at a constant ambient pressure leads to a reduction in the peak pressure of a blast wave. Through this set-up the undisturbed propagation of the blast wave could be compared with the disturbed case. An attenuation of the blast wave was measurable. A further examination through a parameter study led to the following findings.

The measurable attenuation of the blast wave is the more effective the higher the temperature, the nearer the attenuation space to the asset and the more the thickness becomes a value of $D > 1.50$ m. In addition, it was possible to find indications of diffraction or reflection of the blast wave.

With the found set of best parameters a simulation of the basic set-up was performed. It led to a reduction of the peak overpressure of the blast wave with an achieved attenuation of about 35.0%. With respect to the very high overpressure

values of about 63 kPa, this is a reduction which can decrease the damage of a vehicle significantly.

Based on the preceding findings, the basic set-up was improved to a more realistic scenario. Instead of using a huge PETN explosive ordnance, a simplified model of a mortar grenade was developed. This model reflects the properties of a vastly distributed ammunition type, which is a real threat for security forces in many foreign operations. This model allowed to do an advanced simulation of an undisturbed propagation of the blast wave and a propagation through an attenuation space. As result, an attenuation of the blast wave in this more realistic scenario of up to 33.9% was achieved.

As this research has shown, it is theoretically possible to attenuate the blast wave by second, transient medium. In order to further investigate its effectiveness, the next step should be to examine how the shape and a more accurate selection of the gas parameters affect the attenuation. In particular, it must be checked whether a high attenuation can be achieved by new material models such as, for example, air with ionization or an air-dust mixture. An experimental verification of such an attenuation space of heated air would be particularly revealing for a technical realization. Therefore, it is interesting to ask whether the presented concept can be developed with a proper configuration of the second, transient medium to a kind of attenuation shield against blast waves.

REFERENCES

- [1] B. Hillig, A. Ramezani, and H. Rothe, "Verification of the Theoretical Feasibility of Shock Wave Attenuation by Means of a Second Transient Medium for Protecting Vehicles Against Blasts," in *SIMUL 2017 : The Ninth International Conference on Advances in System Simulation (SIMUL 2017)* Athens, Greece. IARIA, Oct. 2017, pp. 68–73, ISBN: 978-1-61208-594-4.
- [2] B. J. Tillotson and M. A. Loudiana, "Method and apparatus for shockwave attenuation," U.S. Patent No. 8,806,945, 2014.
- [3] R. Cobbold, "The Effects of Operations Other Than War-Fighting on the Participants," in *Royal United Services Institute for Defence and Security Studies, Conference, Paris, France, Jun. 2005*, (Keynote Address).
- [4] P. F. Herrly, *The Impact of Peacekeeping and Stability Operations on the Armed Forces*. Heritage Foundation, 2005.
- [5] J. Kress and S. Grogger, "The domestic ied threat," Institute for National Strategic Studies, National Defense University, Washington, Tech. Rep., 2008.
- [6] D. E. Carlucci and S. S. Jacobson, *Ballistics: theory and design of guns and ammunition*. CRC Press, 2013.
- [7] J. A. Zukas, W. Walters, and W. P. Walters, *Explosive effects and applications*. Springer, New York, 1998.
- [8] I. Sochet, *Blast Effects: Physical Properties of Shock Waves*. Springer, 2017.
- [9] W. E. Baker, *Explosions in air*. University of Texas press, 1973.
- [10] O. Igra, J. Falcovitz, L. Houas, and G. Jourdan, "Review of methods to attenuate shock/blast waves," *Progress in Aerospace Sciences*, vol. 58, 2013, pp. 1–35.
- [11] G. L. Gettle and V. H. Homer, "Acoustic/shock wave attenuating assembly," U.S. Patent No. 5,394,786, 1995.

- [12] K. Kitagawa, M. Yokoyama, and M. Yasuhara, "Attenuation of shock wave by porous materials," in *Shock Waves*, Z. Jiang, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1247–1252.
- [13] J. L. Waddell and J. F. Gordon, "Acoustic shock wave attenuating assembly," U.S. Patent No. 8,316,752, 2012.
- [14] E. Lozano and P. Vilem, "Design and testing of blast shields for different blasting applications," *Blasting and Fragmentation*, vol. 9, no. 1, 2015, pp. 1–18.
- [15] M. Sommerfeld, "The unsteadiness of shock waves propagating through gas-particle mixtures," *Experiments in Fluids*, vol. 3, no. 4, 1985, pp. 197–206.
- [16] O. Igra, G. Ben-Dor, F. Aizik, and B. Gelfand, "Experimental and numerical investigation of shock wave attenuation in dust-gas suspensions," in *Shock Waves@ Marseille III*. Springer, 1995, pp. 49–54.
- [17] J. Zukas, *Introduction to hydrocodes*. Elsevier, 2004, vol. 49.
- [18] A. Hamouda and M. Hashmi, "Modelling the impact and penetration events of modern engineering materials: characteristics of computer codes and material models," *Journal of materials processing technology*, vol. 56, no. 1-4, 1996, pp. 847–862.
- [19] D. J. Benson, "Computational methods in lagrangian and eulerian hydrocodes," *Computer methods in Applied mechanics and Engineering*, vol. 99, no. 2-3, 1992, pp. 235–394.
- [20] M. Oevermann, S. Gerber, and F. Behrendt, "Euler-lagrange/dem simulation of wood gasification in a bubbling fluidized bed reactor," *Particuology*, vol. 7, no. 4, 2009, pp. 307–316.
- [21] D. Hicks and L. Liebrock, "Sph hydrocodes can be stabilized with shape-shifting," *Computers & Mathematics with Applications*, vol. 38, no. 5, 1999, pp. 1–16.
- [22] X. Quan, N. Birnbaum, M. Cowler, B. Gerber, R. Clegg, and C. Hayhurst, "Numerical simulation of structural deformation under shock and impact loads using a coupled multi-solver approach," in *Proceedings of 5th Asia-Pacific Conference on Shock and Impact Loads on Structures*, 2003, pp. 12–14.
- [23] A. Ramezani, B. Hillig, and H. Rothe, "A Coupled CFD-FEM Analysis to Simulate Blast Effects on High Security Vehicles Using Modern Hydrocodes," in *SIMUL 2017 : The Ninth International Conference on Advances in System Simulation (SIMUL 2017) Athens, Greece*. IARIA, Oct. 2017, pp. 74–81, ISBN: 978-1-61208-594-4.
- [24] A. AUTODYN, "Autodyn user's manual, release 18.0," 2017.
- [25] M. Johansson, O. P. Larsen, L. Laine, and R. S. AB, "Explosion at an intersection in an urban environment—experiments and analyses," in *Proceedings of the 78th Shock and Vibration Symposium*, Philadelphia, PA, USA, 2007.
- [26] E. W. Lemmon, R. T. Jacobsen, S. G. Penoncello, and D. G. Friend, "Thermodynamic properties of air and mixtures of nitrogen, argon, and oxygen from 60 to 2000 k at pressures to 2000 mpa," *Journal of physical and chemical reference data*, vol. 29, no. 3, 2000, pp. 331–385.
- [27] J. H. Keenan, J. Chao, and J. Kaye, "Gas tables: international version," in *Gas tables: international version*. John Wiley & Sons, 1983.
- [28] D. O. Dusenberry, *Handbook for blast resistant design of buildings*. John Wiley & Sons, 2010.
- [29] I. V. Hogg, *Mortars*. The Crowood Press Ltd, 2001.
- [30] I. V. Hogg, *The Illustrated Encyclopedia of Ammunition*. Chartwell Books, 1985.
- [31] I. V. Hogg, *Ammunition: Small Arms, Grenades, and Projected Munitions*. Greenhill Books/Lionel Leventhal, 1998.
- [32] V. Clancey, "Diagnostic features of explosion damage," in *6th Int. Meeting of Forensic Sciences*, Edinburgh, 1972.
- [33] U. N. R. Commission et al., "Fire dynamics tools (fdts): Quantitative fire hazard analysis methods for the us nuclear regulatory commission fire protection inspection program," NUREG-1805, December, 2004.

Advances in the Development of Combat Helmet Systems

Constructional Contributions for Lightweight Ballistic Composites

Short Paper

Henrik Seeber, Arash Ramezani, Hendrik Rothe
 Chair of Measurement and Information Technology
 Helmut Schmidt University - University of the Federal Armed Forces
 Hamburg, Germany
 Email: henrik.seeber@hsu-hh.de, ramezani@hsu-hh.de, rothe@hsu-hh.de

Abstract—In the military and police sector, the endangerment for the forces has risen over the last decades. Especially police forces are facing new threats due to increased terrorist activity in western European cities and a rising propensity to violence. This development makes it necessary to provide state of the art protection for patrol officers. This includes helmets made of ultra-high molecular weight polyethylene (UHMWPE) to lower the overall weight of equipment and increase the combat value of the forces by providing more comfort and possibilities for attachments. At the moment, these types of helmets are not ready to fulfill the required level of protection against projectiles. The intention of this paper is to give background information about these new threats and to mention first ideas how to tackle the emerging problems of current UHMWPE combat helmets. In addition, early results regarding the process of UHMWPE will be presented.

Keywords - *applied research; fiber-reinforced plastics; optimization; armor systems; ballistic trials; material processing; post-processing.*

I. INTRODUCTION

Combat helmets are a key factor in personal protection for military and police forces. Rapidly changed threats on missions, especially for police patrol officers and soldiers in stabilization missions, prove that protection needs a new ability profile [1]. This work provides background information to specify the problems and first ideas how to solve these. Later on, findings and required settings for pre- and post-processing UHMWPE will be shown. The overall aim is to create a ballistic combat helmet that meets VPAM 3 regulations (the third level of ballistic protection of the European "union of test centers for armored materials and constructions") and the technical directive "System Ballistic Helmet" 5/2010 [2].

II. STRUCTURE OF THE PAPER

Section III gives a brief information about the history of combat helmets and their materials through time. In Section IV, threats for ballistic protection are mentioned. Section V is about the disadvantages of current UHMWPE combat helmets and Section V deals with the advantages of lighter polymer combat helmets. In Section VII, the aim is concretized and explained. In the final section, Section VIII, the steps in the development process of a combat helmet are

described. In addition, findings in the processing and post-processing of UHMWPE are pointed out.

III. HISTORY

Combat helmets have a long tradition. Before the invention of gunpowder, they were used as a protection against blunt trauma and cuts. They were designed to deflect, e.g., a sword, so there was less residual energy on the head. Later on, helmets were mainly worn for pageantry and unit recognition until the First World War began.

A. Combat Helmets in the 20th Century

Due to the massive use and increased lethality of artillery, the German forces introduced the "Steel Helmet Modell 1916" in 1915. All nations introduced nearly the same helmet models in at this time, which were made of basic steel. These helmets were only able to stop the primary threat of that time: fragmenting projectiles of artillery bombs. They were not able to stop bullets because of the available materials. During the Second World War, the U.S. Military introduced the M1 in 1943, which was made out of "Hadfield steel" (see Figure 1). This helmet was used by the German armed forces until the 1990s. On the one hand, problems with this type of steel helmet occurred because the helmets were too heavy and reduced the view, hearing and mobility of the soldier. On the other hand, the helmet provided reliable protection against light and medium fragments. The M1 was followed by a very new generation of combat helmets, which was made of aramid. Aramid was the first synthetic bulletproof material and was invented in the 1960s by DuPont [3].

B. Combat Helmets in the 21st Century

As a replacement of the M1, the US-Military introduced the "Personnel Armor System for Ground Troops" (PASGT), was made of aramid. In addition, they used the new retention system "NOSHA", which provided a better shock absorption and air circulation. The German armed forces introduced this type of helmet system as well, called "Combat Helmet, Ground Forces". This helmet system is barely able to stop fragments and 9mm bullets, but the residual energy is still too high and as such, the helmet cannot provide reliable protection against aforementioned threats. The "Advanced Combat Helmet" (ACH) primary

has a new suspension system to improve protection against blunt impacts. Additionally, the design changed to allow the usage of new equipment, such as ear protection and radio systems. The next stage of development was – again – a totally new material: the ultra-high molecular weight polyethylene fibers (UHMWPE). This material combines low weight (areal density: $\sim 86.18 \text{ N/m}^2$) with high strength (tenacity: $\sim 34 \text{ g/d}$) [4]. With hybridization techniques, the U.S. Military developed a new generation of combat helmets in 2010, the so-called “Future Assault Shell Technology” (FAST). The German armed forces also use FAST with the name “Combat Helmet Special Forces”. Also, FAST helmets were added to the concept “Infantry of the Future” (see Figure 2) [5]. FAST helmets have been designed and developed with special forces and air borne units in mind. The helmet system only provides protection against light fragments and blunt impacts. The retention

system has been upgraded to a multi-pad four-point retention system. This leads to reduced weight and higher comfort for the soldier. In 2012 the “Enhanced Combat Helmet” (ECH) was developed. Due to the hybridization of fiber composites, the helmet is able to withstand ballistic threats. However, the residual energy of a projectile is still too high, which leads to back-face deformation and life-threatening injuries. The Heads-Up helmet system introduced in 2013 aims to protect the entire head and paves the way for a new trend. The current generation of helmets mainly protects the wearer against fragmentation. Combat helmets are developed in order to increase combat effectiveness of the individual soldier. In the future, this should extend to protection from blast and ballistic threats while having fully integrated attachments such as communication systems, ear protection or a head-up display [3].

	M1 Steel Pot 1943	PASGT 1980	ACH 2005	FAST 2010	ECH 2012	HEADS UP 2013	FUTURE 2018
Helmet Design							
Helmet Materials	Hadfield Steel	Aramid Fiber (Kevlar®)	Improved Kevlar®	UHMWPE (Dyneema®) and Carbon Fiber			Future S&T efforts focused on more comprehensive head protection (ballistic/blast/blunt trauma)
Helmet Threats	Fragmentary rounds and .45 M1911 bullet	Fragmentary rounds, 9mm		Fragmentary rounds	Fragmentary rounds, 9mm and specified small arms	Fragmentary rounds at lighter weight, small arms	Address blast as well as ballistic threats
Areal Density	2.2 psf	2.2 psf	2.2 psf	1.8 psf	2.2 psf	1.6 - 1.8 psf	TBD
Tenacity	-----	23 g/d	27 g/d	34 g/d	37 g/d	37 g/d	TBD
Significance	Used in WWII, Korea, and Vietnam	Material revolution: synthetic ballistic material	Changes in design and suspension, improved aramids	Distinctive innovations in design and materials	First helmet with specified frag and small arms protection	Attempting extensive/integrated head protection	Overarching improvements in head protection

Figure 1. Evolution of modern combat helmets



Figure 2. FAST helmet with the concept "Infantry of the Future" [5].

IV. NEW THREATS

A. Changed Threats for Police Forces

Threats for police forces and military troops have changed a lot over the last decades. Police forces face international terrorism, especially in western European cities. More and more, terrorists are professionally trained and equipped with military weapons and gear. Time is the most crucial point in hazardous situations, so patrol officers have to engage first [5]. Only a combination of ballistic vests and ballistic helmets provides the necessary level of protection in such situations. Especially patrol officers are facing unpredictable threats on duty, so their helmets have to provide protection against multiple threats. Apart from this, hits with blunt and sharp weapons, fire and chemicals are common risks for them. The willingness of patrol officers in Baden Württemberg to wear their helmets also in common situations like brawls and skirmishes makes it necessary to provide good shock absorbing attributes against blunt trauma [1][6].

B. New Threats for Military Forces

Military forces are facing changed threats. Statistics of the American operations in Afghanistan and Iraq show that head and neck wounds are increasing. The distribution shows that 30% of all wounds are in the head and neck area (based on injuries/treatments from hospitalization, including persons who died of wounds) [3]. The main threats at patrol missions are improvised explosive devices (IED) and ambushes with assault rifles. Due to the increased use of IEDs, blast associated head injuries, e.g., fragments have increased compared to gunshot wounds. Furthermore, the characteristics of the fragments have changed compared to mortar and artillery shells. This can lead to a different impact behavior. In addition, blunt traumatic injuries have increased because they are linked to blast events. Nevertheless, blunt trauma is also associated with non-combat situations like vehicle crashes, parachute drop accidents or falls. Common blunt trauma threats have an impact velocity of 6.1 m/s, which is equal to a drop of 1.9 m [3]. The primary ballistic threat is caused by assault rifles of type AK-47 (7.62x39-mm) and owing to the increased close combat situation pistols emerging as threats, for example, Makarov (9x18-mm) or Tokarev (7.62x25-mm). Altogether,

the America Department of Defense locates the main threat of infantry weapons at 5.56-mm and 7.62-mm rounds and muzzle velocity from 735 m/s to more than 800 m/s. This matches approximately VPAM 6 to VPAM 7 [1][3].

V. DISADVANTAGES OF ACTUAL UHMWPE COMBAT HELMETS

A. Back-face Deformation

Back-face deformation is one of the main problems of the actual UHMWPE combat helmets. On the one hand, the material has very good attributes against bullet penetration. On the other hand, the energy of the bullet is not well dispersed. In order to understand the back-face deformation, it is vital to understand the behavior of the fiber after an impact. Two waves occur: a transversal and a longitudinal wave (see Figure 3). The longitudinal wave moves along the fiber. During this movement, the fiber is stretched and constricted. The transversal wave moves in the direction of the projectile path. Due to the stretched fiber and the transversal wave the material suffers a deformation in the direction of the projectile. This leads to the so-called back-face deformation, the material indent and the residual energy appeals on the head [4]. For German police helmets, the residual energy has a maximum tolerance limit of 25 Joule [6]. The residual energy could lead to possible head injuries like long linear skull fractures or closed head brain trauma. At the moment, it is unclear whether the injuries occur from the deforming of the helmet onto the head or from acceleration loads transmitted through the helmet padding to the head.

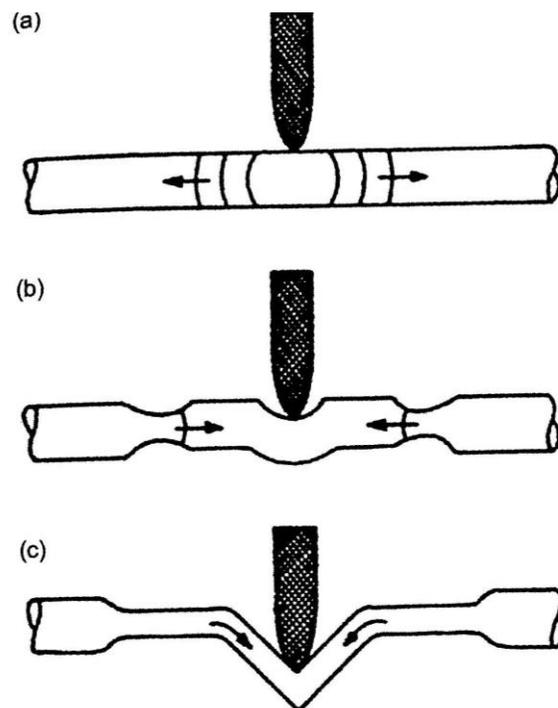


Figure 3. Energy distribution in a fiber impacted by a projectile [4].

In addition, the test methodology with clay to display back-face deformation is not totally linked to head injuries. The human skull behavior in such situations is inadequately represented in the actual test methodology with clay. Especially in the area of back-face deformation there is a lot of potential for necessary improvements [1].

B. Blunt impacts

Moreover, the current generation of UHMWPE combat helmets have deficits with blunt impacts. The current combat helmets can only absorb impacts with a velocity of 3 m/s or 45 J drop energy [3]. As mentioned earlier, common blunt traumas occur with a velocity of 6.1 m/s. In fact, blunt traumas occur especially in non-combat and training situations. Therefore, there needs to be an improvement, because most of the time the wearer of the helmet is in such a situation [1].

VI. COMBAT HELMET AS MODULAR HEADGEAR SYSTEM

Combat helmets will evolve from a device only used for protection to a multi useable platform to increase the survivability and efficiency of the wearer. This includes basic attachments like active ear protection, flashlights or counter weights to provide a stable weight balance. Moreover, the helmet platform can be used to increase the leading ability of the group by adding integrated voice radio, a head up displays with important mission information or health sensors to monitor the group vital functions. This would increase the situational awareness of the group leader and would lead to an overall increase of safety during missions. Furthermore, the combat value of every soldier or policeman can be increased by adding feeder plates for night vision, the ability to wear protective masks against warfare agent or attaching standardized rails like MIL-STD 1913. Of course, the possibilities are limited due to the weight the wearer can handle over the duration of the mission. So, if the combat helmet itself weight as little as possible, there are more possibilities for attachments and this leads to earlier mentioned advantages [1].

VII. AIM OF THE PROJECT

First of all, the aim is to create a UHMWPE combat helmet, which meets the regulations of the police institute of the German police academy in Muenster, this regulation is based on the technical directive "System Ballistic Helmet" (Technische Richtlinie "Gesamtsystem Ballistischer Schutzhelm") from May 2010 [2]. The use test standards are the VPAM "APR 2006" (Allgemeine Prüfrichtlinie 2006; Engl.: general test guideline 2006) and VPAM "HVN 2009" (Durchschusshemmender Helm mit Visier und Nackenschutz 2009; Engl.: bullet-resistant helmet with visor and neck protection 2009).

A. Threats

The main focus of the research is on the ballistic attributes of the helmet. Therefore, the helmet has to provide

protection against soft-core projectiles 9mmx19 fired by small arms and machine pistols. This is comparable to VPAM 3. Furthermore, the aim is to meet the regulations of VPAM 4 to compete with the current generation of titanium helmets and provide a state-of-the-art alternative [5]. The mentioned titan helmet, which is actually used by the state of Baden Wuerttemberg, is the "Hoplit" model by Ulbrichts Witwe GmbH (see Figure 4). As mentioned, earlier protection against blunt trauma is also a challenge for combat helmets. The residual energy has to be lower than 25 Joule [1].

B. Constructional Problems

At the moment, the material has a reliable protection against projectile penetration. Additionally, a possible helmet shape provides a reliable protection. It has to be verified in which areas the protection is effective. Especially near the edges of the calotte, it is possible that the protection efficiency is much lower compared to the central areas. The actual titanium helmets have an efficiency distance of 10 mm to the edges. All in all, these helmets provide an effective protection area of 90% [6]. As with aramid helmets, which have a much lower protective area, the fiber structure of the UHMWPE could also be a crucial point to provide a protection area as big as titanium helmets [5]. Moreover, it is important to determine the optimal shape and configuration in which the UHMWPE tape is used. The goal is to use fibers, which are as long as possible in order to retain the good physical properties. This is difficult to achieve with a spherical object and rectangular tapes.

C. Possible Solutions

The main problem is back-face deformation. The residual energy dispensation of the material is too low in the current configuration. Now three possibilities to increase the dispensation have to be tested. Varying the direction of the material layers may mitigate the deformation. This has to be balanced between penetration and deformation of the material. The best penetration protection is provided when the layers are rotated by 90 degrees. Another idea is to use energy-absorbing materials under the calotte and as helm pads to reduce the residual energy. So, this means to integrate strictly the inlay into the helmet.



Figure 4. "Hoplit F" by Ulbrichts Witwe GmbH [6]

Another possibility is to use two calottes, the first one as a ballistic shell and the second one as a shell to disperse the residual energy and to add absorbing material between the shells. The next step would be to precise the ideas and test their efficiency. After this, a combination of ideas could reduce the residual energy to a value lower than 25 Joule. Finally, the aim is to meet the regulations of VPAM 3 [1].

VIII. COMBAT HELMT DEVELOPMENT

Before the development started, a progress plan was created. The plan is structured as follows: The first step is a detailed research concerning the material involved. This includes a literature research about failure behavior and material properties (a detailed report about this topic is readable [8]) and studies about material behavior during the processing, which is mentioned in this article. The next step will be the production of a prototype, made with different construction methods to find an optimal configuration. Also, the optimal hardening process parameters need to be found. Finally, these prototypes have to be tested on the basis of the VPAM HVN 2009 standards.

A. Used Material

The UHMWPE, which is used for this research is Dyneema® HB26. It is shipped in tape shape and as such, the matrix and fiber have already been combined to a fiber composite material. This material is ideal for body armor due to its low weight (density: 15.9 g/cm^3) and high tenacity of 37 g/d [7]. The tape has a bidirectional fiber structure with four layers of fibers (see Figure 5). This means that the tape has two layers in horizontal and two layers in vertical direction. That is an important fact, because the physical properties depend on the supported fiber length in the significant direction and has to be considered during the construction. Another crucial point is low temperature stability. The operating temperature range must be lower than 80°C , otherwise the material behavior might change [7].



Figure 5. Damaged Dyneema® HB26 tape. Layer-, matrix- and bidirectional fiber structure is visible.

B. Pre-processing

The tape has to be cut into the needed pieces for the construction. Because of the fiber structure of the material, a contactless cutting method is advantageous. Therefore, a laser cutter is used. It uses a CO_2 laser due to the inorganic material, which is processed. Due to the limited working surface of the laser cutter, the tape has to be cut into semi-finished material, which fits into the laser cutter. Therefore, a special fiber scissor is used for a preferably tear free cutting. The laser cutter has three main parameters influencing the cutting result: power, velocity and frequency. Following recommendations of the manufacturer of the laser cutter the settings are chosen as follows: the power is set to 120 W, the frequency is set to 1000 Hz and a 1.5'' lens is used. Now, only the velocity needs to be determined. The cutting results are evaluated in several experiments. Bulging caused by the lead in heat, an optic evaluation of the cut and visible damages are compared at various velocities. Bulging turned out to be an important figure. That is because in lower velocities, the heat input leads to a bead nearly twice as thick as the thickness of the Dyneema® HB26 (original thickness 0.35 mm) (see Table I). The same applies to high velocities. High beads are an indication that the edges have been melted away and causes problems with geometrical accuracy. A velocity of 0.036 m/s seems to be optimal cutting setting to cut one layer of Dyneema® HB26 (see Table II). With this velocity, 0.5 mm of material from edge is melted away. This has to be regarded during the construction of the component. In order to increase productivity, experiments were made with five layers of material taped together being cut at once (see Figure 6). The velocity and focus point of the laser were changed throughout the experiment (see Table III). The best result is achieved with a velocity of 0.011 m/s and a focus point on the first layer leads to an acceptable cut.

TABLE I. COMPARISON OF THE CUT EVALUATION

Velocity [m/s]	Number	Bulging [mm]	Optic Evaluation	Damage
0.027	1	0.65	Bead	Fiber tearing
	2	0.68	Bead	No
0.032	1	0.58	Bead	No
	2	0.59	Bead	No
0.036	1	0.48	Even Edge	No
	2	0.48	Even Edge	No
0.044	1	0.52	Even Edge	No
	2	0.51	Even Edge	No
0.053	1	0.53	Even Edge	No
	2	0.56	Fringed Edge	No
0.062	1	0.59	Fringed Edge	No
	2	0.60	Bead	No
0.071	1	0.68	Bead	No
	2	0.63	Bead	No

TABLE II. OPTIMAL CUTTING PARAMETER OF DYNEEMA® HB26

Parameter	Value
Material	Dyneema® HB26
Power	120 W
Frequency	1000 Hz
Velocity	0.036 m/s

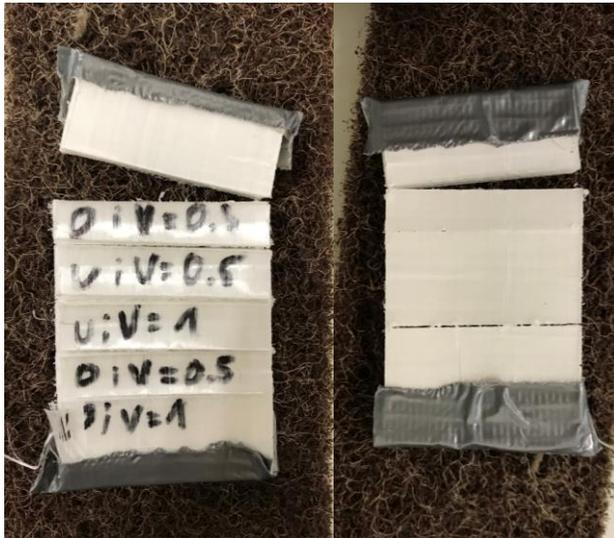


Figure 6. Left: Dyneema HB26 cluster with parameter labeling. Right: Reverse side of the cluster.

TABLE III. EVALUATION OF CUTTING PARAMETERS OF DYNEEMA® HB26 FIVE LAYER CLUSTER

Focus Point	Velocity [m/s]	Successful Cut
Upper Layer	0.011	Yes
Bottom Layer	0.018	No
Bottom Layer	0.036	No
Upper Layer	0.018	No
Upper Layer	0.036	No

C. Post-processing

After the hardening-process, the edges of the plates are often irregular. It is advantageous to achieve a clear edge to prevent fiber tearing and damages. The test plates have a thickness of twenty layers and are hardened. In the evaluation, the velocity, frequency, focus point, lens and repetitions are variable. However, multiple settings are tested and evaluated to the following criteria: geometrical accuracy, damages, bead and cleanliness of the cut. Finally, the settings displayed in Table IV produce an optimal cutting result (see Figure 7).

TABLE IV. OPTIMAL CUTTING PARAMETER OF HARDENED 20 LAYER DYNEEMA® HB26

Parameter	Value
Velocity	0.011 m/s
Frequency	1000 Hz
Focus Point	Upper Layer
Lens	1.5''
Repetitions	5



Figure 7. Hardened 20 layers Dyneema® HB26 plate after post-processing.

IX. CONCLUSION AND FUTURE WORK

There are three main risks for ground forces: the main blast, blunt trauma and ballistic threats. Especially ballistic threats are challenging the UHMWPE helmets because of a high amount of residual energy. This leads to back-face deformation, which can result in live risking head injuries. In addition, this characteristic of injuries appears with blunt traumas. Some of the mentioned ideas could also lower the risk of blunt traumas even if the main challenge is to reduce back-face deformation. Another advantage of lighter helmets is, in addition to more comfort, the ability as a multi role carrier for attachments and a higher acceptance on the part of wearer. This ability could improve the survivability and efficiency of the wearer. Nevertheless, the focus is to reduce back-face deformation to meet the regulations of VPAM 3. Possible ideas are to verify the direction of the layers, using energy-absorbing materials for the helmet inlay or using two decoupled shells with energy absorbing materials in between. Therefore, a combination or balance between the mentioned ideas is necessary. Now, basic information about the material used exists and can be used to develop and build a prototype helmet. The next step is to evaluate how the prototype helmet can be constructed with possible solutions integrated as mentioned in Section VII/C. After that, ballistic tests are necessary to get an overview of the efficiency and how practical the solutions are. Especially the findings relating to the test methodology of the Review of Department of Defense Test Protocols for Combat Helmets [3] could be implemented into the test cycle. Findings mentioned in Section VIII are also usable for every future project with Dyneema® and represent an important first step. However, this should encourage further research of the hardening process to achieve better results in the future. To conclude, back-face deformation is current the main problem of UHMWPE helmets due to the residual energy transmitted through the inlay. As mentioned in Section V, also, the test methodology, to investigate the relation between back-face deformation and head injuries, has to be beheld and then maybe adjusted.

REFERENCES

- [1] H. Seeber, A. Ramezani, and H. Rothe, "Future Tasks of Lightweight-Polymer Combat Helmets", The Ninth International Conference on Advances in System Simulation (SIMUL2017) IARIA, Oct. 2017.
- [2] Police Institute of the German Police Academy in Muenster, Technical Directive "System Ballistic Helmet" (Technische Richtlinie "Gesamtsystem Ballistischer Schutzhelm"). Muenster, pp. 10-16, 5/2010.
- [3] Committee on Review of Test Protocols Used by the DOD to Test Combat Helmets, "Review of Department of Defense Test Protocols for Combat Helmets", Washington: The National Academies Press, pp. 8-24, 2014.
- [4] I. G. Crouch, L. Arnold, A. Pierlot, and H. Billion, "Fibres, textiles and protective apparel" in "The Science of Armour Materials". Amsterdam: Elsevier, p. 271, 2017.
- [5] J. Weisswange, Protection for the "Main"-Weapon, Modern Combat Helmets (Schutz für die "Haupt"-Waffe, Moderne Gefechtshelme). [Online]. Available from: [http://www.esut.de/fileadmin/user_upload/EST916 Schutz_fuer_die_Hauptwaffe_Moderne_Gefechtshelme.pdf](http://www.esut.de/fileadmin/user_upload/EST916_Schutz_fuer_die_Hauptwaffe_Moderne_Gefechtshelme.pdf) [08.01.2017].
- [6] Polizei Praxis. Ballistic Helmets for Patrol Officers (Ballistischer Schutzhelm für Streifenpolizisten). [Online]. Available from: <http://www.polizeipraxis.de/themen/arbeitsschutz-in-der-polizei/detailansicht-arbeitsschutz-in-der-polizei/artikel/ballistischer-schutzhelm-fuer-streifenpolizisten.html> [08.01.2017].
- [7] T. Tam and A. Bhatnagar, "High-performance ballistic fibers and tapes" in "Lightweight Ballistic Composites: Military and Law-Enforcement Applications", Second Edition. Amsterdam: Elsevier, pp. 1-30 and 210-214, 2016.
- [8] H. Seeber, "Contributions to the development of combat helmets made of fiber composite materials for the improvement of ballistic attributes", in: Bachelor degree thesis at Helmut Schmidt University- University of the Federal Armed Forces Hamburg, Dec. 2017.

Modeling and Simulation of Enzyme-Catalysed Substrate Conversion in a Microbioreactor

Linas Petkevičius and Romas Baronas

Institute of Computer Science, Vilnius University,
Didlaukio st. 47, Vilnius, Lithuania

Email: {linas.petkevicius, romas.baronas}@mif.vu.lt

Abstract—Mathematical modeling and numerical simulation of non-linear mathematical model for digital simulation of an enzyme loaded porous microreactor is investigated in this paper. The model is based on a system of reaction-diffusion equations, containing a non-linear term related to the Michaelis-Menten kinetics, and involves three regions: the enzyme microreactor where the enzyme reaction as well as mass transport by diffusion take place, a diffusion limiting region (the Nernst layer), where only the mass transport by diffusion takes place, and a convective region, where the constant analyte concentration is maintained. The digital simulation was carried out using the finite difference technique. Assuming well-stirred conditions, impact of the diffusion modulus and the Biot number as well as porosity on transient effectiveness factor and yield factor has been numerically investigated. The simulation results showed that the Nernst layer must be taken into consideration when modelling micro-size bioreactors. The computational simulations showed that the yield factor and the effectiveness factors have different properties to describe system productivity.

Keywords—reaction-diffusion; Michaelis - Menten kinetics; microbioreactor; porosity; CSTR.

I. INTRODUCTION

This paper is an extension of the work originally reported in The Ninth International Conference on Advances in System Simulation [1].

Continuous-flow stirred tank reactors (CSTR) are common in chemical industries [2][3]. Although a stirred tank is a usual construction of industrial enzyme reactors, the effectiveness and optimal construction of CSTR remain open to study [4][5]. Specifically, further research is needed due to the application of the immobilized enzymes, such as biocatalysts, on a manufacturing scale that requires to use the reactors of different types, including CSTR [6][7][8].

The CSTR often refers to the model used to estimate the operation parameters when using a continuous agitated-tank reactor to reach a specified output [7]. In the last few decades, immobilized enzyme reactor models have evolved significantly with a wide range of applications in food industry [9], waste cleaning [10], immobilization of microbial cells [11][12]. The main advantages of immobilized cells over non-immobilized cells can be enumerated: (1) incapsulation not only keep enzyme from wash out, but also protect it during the changes of various conditions, such as pH, temperature, and toxic compounds; (2) cells do not stick up to bioreactors

or other materials within the reactor; (3) reusability; and (4) control over reactions that are allowed, shell can be formatted to avoid the diffusion of all material to microreactor. Rapid progress was noticed in integrating microfluidic reactors and biocatalytic reactions [13]. The combination of miniaturized technologies and microfluidics allows to increase the efficiency of bioprocess. In the recent review, Jansen concludes that end-to-end fabrication and on-demand manufacturing based on microreactors have growing industrial interest, which leads to development and progress [14]. However, coupling microreactors and biocatalysis is highly complex, requiring an integrated approach addressing biocatalyst features, reaction kinetics, mass transfer and microreactors geometry [13][15].

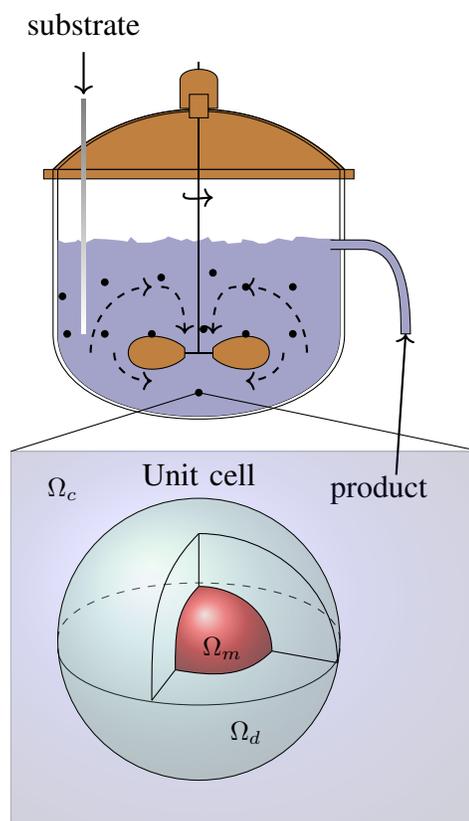


Figure 1. Continuous stirred tank reactor with enzyme-loaded microreactors (pellets) and a zoomed unit cell to be modelled.

Mathematical models have been widely used to investigate the kinetic peculiarities of the enzyme microreactors [8][13]. Models coupling the enzyme-catalysed reaction with the diffusion in enzyme microreactors are usually used. Since containing catalytic particles, the analyte in CSTR is well-stirred and set in powerful motion, the mass transport by diffusion outside the microreactors is usually neglected [15][16]. Mixing within CSTR creates the Nernst layer while increasing stirrer speed can reduce the thickness of Nernst layer. However, due to the inner force of attraction, the zero thickness of the diffusion shell (layer) can not be achieved [17]. We consider an array of identical spherical microreactors placed in the CSTR shown in Figure 1 [7], where area Ω_m denotes a microreactor, Ω_d denotes surrounding diffusion shell and Ω_c is a convective region.

The goal of this work was to investigate the dependencies of the internal and external diffusion limitations on the yield and the effectiveness factors, modelled by reaction-diffusion equations, containing a non-linear term related to Michaelis-Menten kinetics [7][8][18]. The model involves three regions: the enzyme microreactor, where the enzyme reaction, as well as the mass transport by diffusion take place, a diffusion limiting region, where only the mass transport by diffusion takes place, and a convective region, where the constant analyte concentration is maintained. In practice, the laboratory experiments aimed to be performed with as small as possible particle layer size to avoid process limitations caused by diffusion of the reacting solution in the catalyst pore system. There is a practice of performing experiments under the fluid-dynamic conditions in order to eliminate the lack of mass transfer limitations outside the particles. However, it is not always possible to reduce the particle size as much as it is required for avoiding diffusion limitations.

It has been shown that for first order chemical reactions, the transient effectiveness factor depends not only on the Thiele modulus but also on the adsorption capacity of the system, the fact that should be taken into account when kinetic experiments are designed. However, the type of immobilization can raise the diffusion difference within the reactor and the Nernst layer. The porous materials as well as nanoparticles immobilized in microreactors, can be considered porous with respect to the Nernst layer. This paper also analyses how the porosity impacts effectiveness factor as well as the yield factor. Due to a strong non-linearity of the reaction term, the computer simulation was carried out using the finite difference technique [19].

The rest of the paper is organised as follows: in Section II, the mathematical model and microbioreactor characteristics are described; Section III formulates a dimensionless model and determines the main parameters of the bioreactor; Section IV describes the numerical model and the simulator; in Section V, the results of numerical experiments are presented, and conclusions close the article.

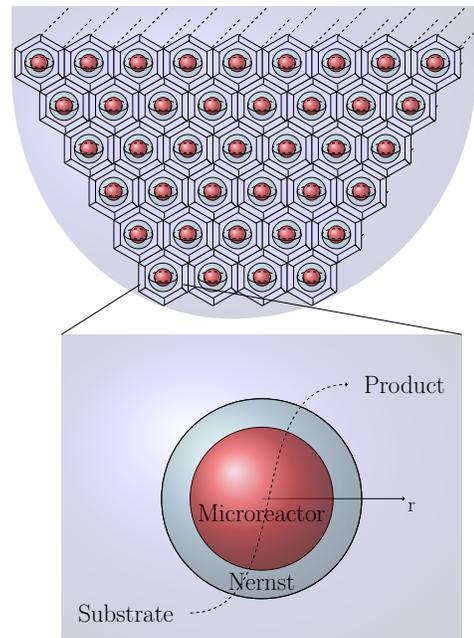
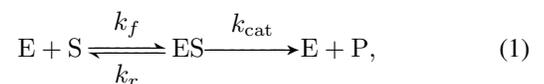


Figure 2. Principal structure of the unit cell consisting of a microbioreactor and the Nernst diffusion shell in homogeneously distributed volume.

II. MATHEMATICAL MODEL

We consider an array of identical spherical microreactors placed in a continuous ideally stirred-tank reactor [7]. Assuming a uniform distribution of the microreactors (see Figure 2) in the tank and a comparably great distance between adjacent microreactors and fact that microreactor volume V_m is significantly smaller than the tank volume V , the spherical unit cell was modelled by an enzyme-loaded microreactor (pellet) and a surrounding diffusion shell (the Nernst diffusion layer). The principal structure of the tank containing uniformly distributed microreactors and a unit cell are presented in Figure 1, where Ω_m denotes a microreactor (MR), Ω_d stands for the diffusion shell and Ω_c is a convective region.

In the enzyme-loaded MR layer we consider the enzyme-catalyzed reaction



the substrate (S) combines reversibly with an enzyme (E) to form a complex (ES). The complex then dissociates into the product (P) and the enzyme is regenerated [20][21]. The k_f , k_r and k_{cat} are forward rate, reverse rate and catalytic rate constant, respectively.

Assuming the steady-state approximation, the concentration of the intermediate complex (ES) does not change and may be neglected when modelling the biochemical behaviour of the microreactor [7][21][22]. In the resulting scheme, the substrate (S) is enzymatically converted to the product (P),



A. Governing Equations

Considering the symmetrical geometry of the microreactor and homogenised distribution of the immobilized enzyme inside the porous microreactor, the mathematical model can be described in one-dimensional domain using the radial distance [7].

Coupling enzymatic reaction in the microreactor (region Ω_m) with the one-dimensional-in-space diffusion, described by Fick's second law, and considering the steady-state of a system (2) leads to the following governing equations of the reaction-diffusion type ($0 < r < R_0$):

$$D_{S,m} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dS_m}{dr} \right) = \frac{V_{\max} S_m}{K_M + S_m}, \quad (3a)$$

$$D_{P,m} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dP_m}{dr} \right) = - \frac{V_{\max} S_m}{K_M + S_m}, \quad (3b)$$

r stands for space variable, $S_m = S_m(r)$ and $P_m = P_m(r)$ are the concentrations of the substrate and the reaction product in the microreactor, respectively, R_0 is the radius of the microreactor, $V_{\max} = k_{cat} E_0$ is the maximal enzymatic rate and $K_M = (k_r + k_{cat})/k_f$ is the Michaelis constant. The volumetric reaction as a function from concentration is marked as $V(S) = V_{\max} S_m / (K_M + S_m)$.

In the Nernst diffusion layer Ω_d only the mass transport by diffusion takes place ($R_0 < r < R_1$):

$$D_{S,d} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dS_d}{dr} \right) = 0, \quad (4a)$$

$$D_{P,d} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dP_d}{dr} \right) = 0, \quad (4b)$$

$S_d = S_d(r)$ and $P_d = P_d(r)$ are the concentrations of the substrate and the reaction product in the diffusion shell. Respectively, $D_{S,d}$ and $D_{P,d}$ are the diffusion coefficients of the materials in the bulk solution and R_1 is the radius of the unit cell.

Microbioreactors are microparticles immobilized with the enzyme, so the core of the particle is porous. Due to the porosity of microbioreactor diffusion coefficients $D_{S,m}$ and $D_{P,m}$ can be expressed through diffusion coefficients for bulk solution [23],

$$D_{S,m} = \frac{D_{S,d} \epsilon_t}{\tau} \quad (5a)$$

$$D_{P,m} = \frac{D_{P,d} \epsilon_t}{\tau}, \quad (5b)$$

ϵ_t is porosity (void fraction space in a material) and τ is tortuosity.

B. Boundary Conditions

Fluxes of the substrate and the product through the stagnant external diffusion shell is assumed to be equal to the corresponding fluxes entering the surface of the microreactor,

$$\frac{\epsilon_t}{\tau} \frac{dS_m}{dr} \Big|_{r=R_0} = \frac{dS_d}{dr} \Big|_{r=R_0}, \quad (6a)$$

$$\frac{\epsilon_t}{\tau} \frac{dP_m}{dr} \Big|_{r=R_0} = \frac{dP_d}{dr} \Big|_{r=R_0}. \quad (6b)$$

The formal partition coefficient ϕ is used to describe the specificity of concentration distribution of the compounds between two neighboring regions [7][24],

$$S_m(R_0) = \phi S_d(R_0), \quad (7a)$$

$$P_m(R_0) = \phi P_d(R_0). \quad (7b)$$

The partition within the reactor caused by porosity, so we assume that the porosity and partition coefficient is in a linear relationship:

$$\phi = \epsilon_t \quad (8)$$

Due to the symmetry of the microreactor, the zero-flux boundary conditions are defined for the center of the microreactor ($r = 0$),

$$\frac{D_{S,d} \epsilon_t}{\tau} \frac{dS_m}{dr} \Big|_{r=0} = 0, \quad (9a)$$

$$\frac{D_{P,d} \epsilon_t}{\tau} \frac{dP_m}{dr} \Big|_{r=0} = 0. \quad (9b)$$

According to the Nernst approach, the shell of thickness $\nu = R_1 - R_0$ remains unchanged with time [17][19]. Away from it, the solution is in motion and is uniform in concentration. Due to the continuous injection of the substrate into the stirred tank and washing off the product, the concentration in the convective region remains unchanged:

$$S_d(R_1) = S_0, \quad (10a)$$

$$P_d(R_1) = 0. \quad (10b)$$

The thickness ν of the Nernst diffusion shell depends on the nature and stirring up of the buffer solution. Usually, more intensive stirring corresponds to the thinner diffusion layer (shell).

C. Microbioreactor characteristics

In many industrial processes, especially in the production of low-value added products like biopesticides, bio-fertilizers, bio-surfactants ect. [10], it is important to continuously improve the yield and/or productivity of the reactor system [8]. The productivity is important, since this ensures an efficient utilization of the production capacity, i.e., the bioreactors.

The yield of the desired product on the substrate is one of the most important criteria for design and optimization of bioreactors. The economic feasibility of the process is expressed by the yield factor as the ratio of product formation rate and the substrate uptake rate [7][8].

The bioreactor construction is efficient enough when the product emission is relatively large with given substrate amount used. The product emission rate $\bar{E}_{P,O}$ can be calculated by an integration of the product flux over the outer surface of the diffusion shell [8],

$$\begin{aligned}\bar{E}_{P,O} &= -\int_0^\pi \int_0^{2\pi} D_{P,d} \frac{dP_d}{dr} \Big|_{r=R_1} R_1^2 \sin(\theta) d\varphi d\theta \\ &= -4\pi R_1^2 D_{P,d} \frac{dP_d}{dr} \Big|_{r=R_1}.\end{aligned}\quad (11)$$

The product emission rate has the units of mol/s.

The substrate consumption rate \bar{C}_S over the whole microreactor can be calculated as follows:

$$\begin{aligned}\bar{C}_S &= \int_0^{R_0} \int_0^{2\pi} \int_0^\pi V(S) \sin(\theta) d\theta d\varphi r^2 dr \\ &= \int_0^{R_0} \int_0^{2\pi} 2V(S) d\varphi r^2 dr \\ &= 4\pi \int_0^{R_0} V(S) r^2 dr. \\ &= 4\pi \int_0^{R_0} \frac{V_{\max} S_m(r)}{K_M + S_m(r)} r^2 dr.\end{aligned}\quad (12)$$

The substrate emission rate also have the units of mol/s.

The yield factor γ for the microreactor system, as well as for the entire tank reactor shown in Figure 1, can be defined by the ratio of the product emission rate to the substrate consumption rate,

$$\gamma = \frac{\bar{E}_{P,O}}{\bar{C}_S}.\quad (13)$$

The yield factor is characterised by an actual rate within the reactor in contrast to the actual product formation rate.

The effectiveness factors characterise the interaction between action in microreactor in porous catalytic pellets and microreactors when particles are solid [25], and often used in the biochemical engineering [26][27]. The effectiveness factors are usually defined in terms of the stationary mode of biocatalytic systems [25][27][28]. The effectiveness factor η can be calculated [29]:

$$\begin{aligned}\eta &= \left(\frac{\bar{C}_S}{(4\pi R_0^3/3)} \right) / V(S_0) \\ &= \frac{3 \int_0^{R_0} V(S) r^2 dr}{R_0^3 V(S_0)} \\ &= \frac{3 \int_0^{R_0} \frac{V_{\max} S_m(r)}{K_M + S_m(r)} r^2 dr}{R_0^3 \frac{V_{\max} S_0}{K_M + S_0}}.\end{aligned}\quad (14)$$

Summarising definition (14), the effectiveness factor η can be defined also as the ratio of the average reaction rate actually observed in the MR to ideal rate evaluated at the bulk concentrations of the substrate [29][30].

III. DIMENSIONLESS MODEL

In order to define the main governing parameters of the two compartment model (3)-(10), the dimensional variable r

and unknown concentrations $S_m(r), P_m(r), S_d(r), P_d(r)$ are replaced with the following dimensionless parameters:

$$\begin{aligned}\tilde{r} &= \frac{r}{R_0}, \quad \tilde{S}_m = \frac{S_m}{K_M} \\ \tilde{P}_m &= \frac{P_m}{K_M}, \quad \tilde{S}_d = \frac{S_d}{K_M}, \quad \tilde{P}_d = \frac{P_d}{K_M},\end{aligned}\quad (15)$$

\tilde{r} is the dimensionless distance from the microreactor center and $\tilde{S}_m(\tilde{r}), \tilde{P}_m(\tilde{r}), \tilde{S}_d(\tilde{r}), \tilde{P}_d(\tilde{r})$ are the dimensionless concentrations. Having defined dimensionless variables and unknowns, the following dimensionless parameters characterize the domain geometry and the substrate concentration in the bulk:

$$\tilde{\nu} = \frac{\nu}{R_0}, \quad \tilde{S}_0 = \frac{S_0}{K_M}, \quad D = \frac{\tau}{\epsilon_t}\quad (16)$$

$\tilde{\nu}$ is the dimensionless thickness of the Nernst diffusion layer, \tilde{S}_0 is the dimensionless substrate concentration in the bulk solution, D is the dimensionless diffusion coefficient. The dimensionless thickness of the microreactor equals one.

A. Model description

The governing equations (3) in the dimensionless coordinates are expressed as follows ($0 < \tilde{r} < 1$):

$$\frac{1}{\tilde{r}^2} \frac{d}{d\tilde{r}} \left(\tilde{r}^2 \frac{d\tilde{S}_m}{d\tilde{r}} \right) - \sigma^2 \frac{\tilde{S}_m}{1 + \tilde{S}_m} = 0,\quad (17a)$$

$$\frac{1}{\tilde{r}^2} \frac{d}{d\tilde{r}} \left(\tilde{r}^2 \frac{d\tilde{P}_m}{d\tilde{r}} \right) + \sigma^2 \frac{\tilde{S}_m}{1 + \tilde{S}_m} = 0,\quad (17b)$$

σ is the Thiele modulus or the Damköhler number [8][31][32] defined as:

$$\sigma^2 = \frac{V_{\max} R_0^2}{K_M D_{S,m}} = \frac{V_{\max} R_0^2 \tau}{K_M D_{S,d} \epsilon_t}.\quad (18)$$

The governing equations (4) take the following form ($1 < \tilde{r} < 1 + \tilde{\nu}$):

$$D \frac{1}{\tilde{r}^2} \frac{d}{d\tilde{r}} \left(\tilde{r}^2 \frac{d\tilde{S}_d}{d\tilde{r}} \right) = 0,\quad (19a)$$

$$D \frac{1}{\tilde{r}^2} \frac{d}{d\tilde{r}} \left(\tilde{r}^2 \frac{d\tilde{P}_d}{d\tilde{r}} \right) = 0.\quad (19b)$$

The matching conditions (6)-(10) become:

$$\frac{d\tilde{S}_m}{d\tilde{r}} \Big|_{\tilde{r}=1} = D \frac{d\tilde{S}_d}{d\tilde{r}} \Big|_{\tilde{r}=1}\quad (20a)$$

$$\frac{d\tilde{P}_m}{d\tilde{r}} \Big|_{\tilde{r}=1} = D \frac{d\tilde{P}_d}{d\tilde{r}} \Big|_{\tilde{r}=1},\quad (20b)$$

$$\tilde{S}_m(1) = \phi \tilde{S}_d(1),\quad (21a)$$

$$\tilde{P}_m(1) = \phi \tilde{P}_d(1),\quad (21b)$$

$$\frac{d\tilde{S}_m}{d\tilde{r}} \Big|_{\tilde{r}=0} = 0,\quad (22a)$$

$$\frac{d\tilde{P}_m}{d\tilde{r}} \Big|_{\tilde{r}=0} = 0,\quad (22b)$$

$$\tilde{S}_d(1 + \tilde{\nu}) = \tilde{S}_0, \quad (23a)$$

$$\tilde{P}_d(1 + \tilde{\nu}) = 0. \quad (23b)$$

B. Model reduction

The second order homogeneous differential equations (19) can be easily integrated ($1 < \tilde{r} < 1 + \tilde{\nu}$) [19],

$$\tilde{S}_d(\tilde{r}) = \frac{c_1}{\tilde{r}} + c_2, \quad \tilde{P}_d(\tilde{r}) = \frac{d_1}{\tilde{r}} + d_2, \quad (24)$$

c_1, c_2, d_1 and d_2 are constants of integration. By evaluating the integration constants from the boundary conditions (20) and (23), we get the following solution to the system (19):

$$\tilde{S}_d(\tilde{r}) = \tilde{S}_0 - \frac{\epsilon_t}{\tau} \left(\frac{1}{\tilde{r}} - \frac{1}{1 + \tilde{\nu}} \right) \frac{d\tilde{S}_m}{d\tilde{r}} \Big|_{r=1}, \quad (25a)$$

$$\tilde{P}_d(\tilde{r}) = -\frac{\epsilon_t}{\tau} \left(\frac{1}{\tilde{r}} - \frac{1}{1 + \tilde{\nu}} \right) \frac{d\tilde{P}_m}{d\tilde{r}} \Big|_{r=1}. \quad (25b)$$

After applying the obtained expressions (25), the matching conditions (21) can be expressed as the flux boundary conditions,

$$\frac{d\tilde{S}_m}{d\tilde{r}} \Big|_{\tilde{r}=1} = \beta_S \left(\tilde{S}_0 - \frac{\tilde{S}_m(1)}{\phi} \right), \quad (26a)$$

$$\frac{d\tilde{P}_m}{d\tilde{r}} \Big|_{\tilde{r}=1} = \beta_P \left(-\frac{\tilde{P}_m(1)}{\phi} \right), \quad (26b)$$

β_S and β_P are the Biot numbers for the substrate and the product [33][34], respectively.

$$\begin{aligned} \beta_S &= \frac{D_{S,d}}{D_{S,d}\epsilon_t} \frac{1 + \tilde{\nu}}{\tilde{\nu}} = \frac{D_{S,d}R_1}{D_{S,d}\epsilon_t(R_1 - R_0)} \\ \beta_P &= \frac{D_{P,d}}{D_{P,d}\epsilon_t} \frac{1 + \tilde{\nu}}{\tilde{\nu}} = \frac{D_{P,d}R_1}{D_{P,d}\epsilon_t(R_1 - R_0)} \\ \beta &= \beta_S = \beta_P = \frac{\tau}{\epsilon_t} \frac{1 + \tilde{\nu}}{\tilde{\nu}} = D \frac{R_1}{(R_1 - R_0)}. \end{aligned} \quad (27)$$

The boundary value problem (17)-(23) has been reduced to the boundary value problem described by the governing equations (17), the boundary condition (22) and the flux boundary condition (26).

C. Governing parameters

The initial set of model parameters has been reduced to the following aggregate dimensionless parameters: $\tilde{\nu}$ is the relative thickness of the Nernst diffusion layer, σ is the diffusion module, β is the Biot number, \tilde{S}_0 is the substrate concentration in the bulk, ϕ is the formal partition coefficient.

The dimensionless factor σ^2 essentially compares the rate of enzyme reaction (V_{\max}/K_M) with the diffusion through the enzyme-loaded microreactor ($D_{S,m}/R_0^2$). If $\sigma^2 \ll 1$, the

enzyme kinetics controls the bioreactor action. The action is under diffusion control when $\sigma^2 \gg 1$.

The Biot number indicates the internal mass transfer resistance to the external one [33][34]. When the Biot number is small, the effect of the external diffusion is the most marked. As the Biot number increases, the effect of the external diffusion becomes less important.

The diffusion module and the Biot number are widely used in analysis and design of different bioreactors [34]. The experiment conducted by Kont et al. [11] proved the external mass-transfer limitations to be negligible for the Biot number greater than one using the first order kinetics model of CSTR and packed-bed reactors (PBR), which corresponds condition (10). Typically, designers seek for bioreactors acting in the reaction-limited regime, since in this case reaction and diffusion occur on different time scales [35].

D. Limiting cases

Analytical solutions of product and substrate concentrations can be found for the limiting cases when volumetric reactions become zero or first kinetics [7][33]. Analytical solutions have practical value, the models when diffusion is not fixed and depends on porosity and tortuosity are rarely analyzed (see equation (5)).

Zero order kinetics: When the substrate concentration S_0 is very high compared to the Michaelis constant K_M ($S_0 \gg K_M$), the reaction term reduces to the zero order reaction rate V_{\max} , then volumetric reactions becomes $V(S) \approx V_{\max}$.

The solution of substrate and product concentrations with the boundary conditions (6)-(10) gives the following expressions:

$$S_m(r) = \phi S_0 - \frac{V_{\max}\tau}{6D_{S,d}\epsilon_t} \left(R_0^2 - r^2 + \frac{2\phi R_0^2}{\beta} \right) \quad (28)$$

$$P_m(r) = \frac{V_{\max}\tau}{6D_{P,d}\epsilon_t} \left(R_0^2 - r^2 + \frac{2\phi R_0^2}{\beta} \right) \quad (29)$$

$$S_d(r) = S_0 - \frac{V_{\max}R_0^3}{3D_{S,d}} \left(\frac{1}{r} - \frac{1}{R_1} \right) \quad (30)$$

$$P_d(r) = \frac{V_{\max}R_0^3}{3D_{P,d}} \left(\frac{1}{r} - \frac{1}{R_1} \right) \quad (31)$$

Corollary 1. *The yield and effectiveness factors are equal to one for zero order kinetics.*

$$\begin{aligned} \bar{E}_{P,O}^0 &= -4\pi R_1^2 D_{P,d} \frac{dP_d}{dr} \Big|_{r=R_1} \\ &= -4\pi R_1^2 D_{P,d} \frac{(-1)}{3} R_0^3 \frac{V_{\max}}{D_{P,d}R_1^2} \\ &= \frac{4}{3}\pi R_0^3 V_{\max} \\ \bar{C}_S^0 &= \frac{4}{3}\pi R_0^3 V_{\max} \\ \gamma^0 &= \frac{\bar{E}_{P,O}^0}{\bar{C}_S^0} = 1 \end{aligned} \quad (32)$$

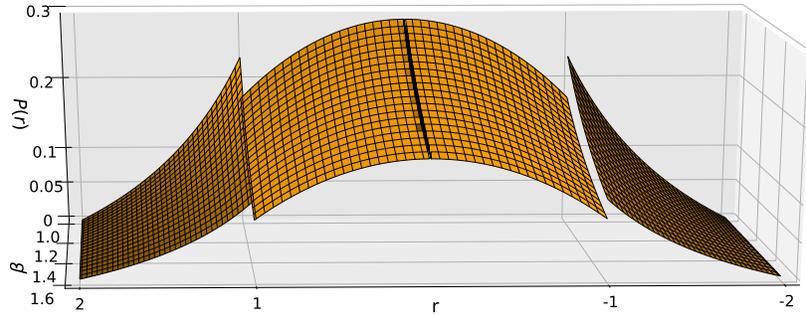


Figure 3. Concentration profiles of the product concentration in the microreactor (simulated at the substrate concentration $\tilde{S}_0 = 1$ porosity $\epsilon_t = 0.75$ and different values of the Biot number $\beta \in [1, 1.6]$, the other parameters are as defined in (40)).

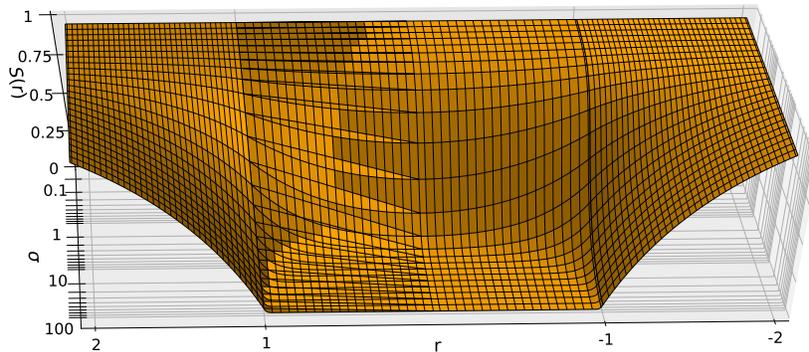


Figure 4. Concentration profiles of the substrate concentration in the microreactor (simulated at the substrate concentration $\tilde{S}_0 = 1$, porosity $\epsilon_t = 1$ and different values of the Thiele module $\sigma \in [10^{-1}, 10^2]$, the other parameters are as defined in (40)).

$$\eta^0 = \frac{3 \int_0^{R_0} V_{\max} r^2 dr}{R_0^3 V_{\max}} = 1 \quad (33)$$

equations (3) reduces to the first order reaction rate, then volumetric reactions becomes $V(S) \approx V_{\max} S / K_M$.

The zero order kinetics rate does not depend on the concentration profiles, so, naturally, such type of system is efficient.

The solution of substrate and product concentrations with the boundary conditions (6)-(10) at first order kinetics gives the following expressions:

First order kinetics: At such low concentration of the substrate as $S_0 \ll K_M$, the non-linear reaction rate in

$$S_m(r) = \frac{\phi S_0 R_1 R_0 \tau / \epsilon \sinh(mr)}{r(\tau / \epsilon R_1 \sinh(\sigma) + \phi(R_1 - R_0)(\sigma \cosh(\sigma) - \sinh(\sigma)))}, \quad (34)$$

$$S_d(r) = \frac{S_0 R_1 (\tau / \epsilon r \sinh(\sigma) + \phi(r - R_0)(\sigma \cosh(\sigma) - \sinh(\sigma)))}{r(\tau / \epsilon R_1 \sinh(\sigma) + \phi(R_1 - R_0)(\sigma \cosh(\sigma) - \sinh(\sigma))} \quad (35)$$

$0 \leq r \leq R_0$.

$$P_m(r) = -\left(\phi S_0 (R_1 \phi \sigma \cosh(\sigma)r - R_0 \phi \sigma \cosh(\sigma)r - \sinh(mr)R_0 R_1 D^* \tau / \epsilon - R_1 \phi \sinh(\sigma)r + R_1 \sinh(\sigma)D^* \tau / \epsilon r + R_0 \phi \sinh(\sigma)r) \right) / \left(m^2 r (\tau / \epsilon R_1 \sinh(\sigma) + \phi(R_1 - R_0)(\sigma \cosh(\sigma) - \sinh(\sigma))) D^* \right), \quad (36)$$

$$P_d(r) = -\frac{\phi S_0 R_0 (R_1 - r)(\sigma \cosh(\sigma) - \sinh(\sigma))}{m^2 r (\tau / \epsilon R_1 \sinh(\sigma) + \phi(R_1 - R_0)(\sigma \cosh(\sigma) - \sinh(\sigma))) D^*} \quad (37)$$

$R_0 \leq r \leq R_1$.

Corollary 2. The yield and effectiveness factors have the following form at first order kinetics.

$$\gamma^1 = \sigma^2, \quad (38)$$

$$\eta^1 = \frac{3\beta\phi(\sigma \coth \sigma - 1)}{\sigma^2(\beta + \phi(\sigma \coth \sigma - 1))}. \quad (39)$$

One can see that the effectiveness coefficient lies in the interval $[0, 1]$, while yield can grow. However, in practical applications $\sigma \ll 1$.

IV. DIGITAL SIMULATION OF EXPERIMENTS

The non-linearity of the governing equations prevents us from solving the boundary value problem (17)-(23) analytically, hence the numerical model was constructed and solved using finite difference technique [19]. An explicit scheme was used; however, due to Michaelis-Menten non-linearity, further construction of equations was used:

$$D_{C,m} \cdot \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dC_m^n}{dr} \right) = \pm \frac{V_{\max} C_m^n}{K_M + C_m^{n-1}},$$

$C = S, P$. Tridiagonal matrix was constructed from the equations. In the numerical simulation, the scheme was run until the following loss became very small:

$$\mathcal{L} = \|S^n - S^{n-1}\|_{l^2} + \|P^n - P^{n-1}\|_{l^2} < \epsilon,$$

decay rate value $\epsilon = 10^{-14}$ was used over l^2 norm. An explicit scheme of finite difference was built on a uniform discrete grid with 128 points in space direction [18]. The simulator has been programmed by the authors in C++ language [36].

The numerical solution of the mathematical model (17)-(23) was validated by using the exact analytical solutions known for very special cases of the model parameters [7][18][31][34] see equations (28)-(37).

V. RESULTS AND DISCUSSION

To investigate the effects of the geometry and catalytic activity of the microreactor, the reactor action was simulated and the yield and the effectiveness factor was calculated for very different values of the Biot number β , the Thiele module σ , the substrate dimensionless concentration S_0 , the porosity ϵ_t .

A. Concentration profiles

Figure 3 shows the profiles of the product concentration P calculated from the microreactor model (17)-(23) changing the Biot number β and the porosity equal $\epsilon_t = 0.75$. Figure 4 shows the profiles of the substrate concentration S changing the Thiele module σ and the porosity $\epsilon_t = 1$. As both figures demonstrate, the following parameters of the model remain unchanged:

$$D_{S,d} = D_{P,d}, \quad \beta = 1, \quad \tilde{\nu} = 1. \quad (40)$$

One can be seen in Figure 4, where low Thiele modulus values $\sigma < 1$, which indicates that the species pass the Nerst

diffusion layer fast, allow the concentrations approach the straight line because of linearity of governing equations in the area $\tilde{r} \in (1, 1 + \tilde{\nu})$. On the other hand, high Thiele modulus values ($\sigma \geq 100$) (see Figure 3) lead to significant differences in concentration distribution across the outer boundary of the microreactor. Both the radius of dimensionless microreactor and the Nerst layer thickness is equal to one.

B. The impact of the Biot number, Thiele module and substrate concentration

To investigate the dependence of the yield factor γ , the effectiveness factor η on the Biot number β , changing the Biot number in a range of $[1, 10^2]$, the factor γ was calculated at different values of the Thiele module σ in a range of $[10^{-1}, 10^2]$ and the substrate concentration \tilde{S}_0 was calculated in a range of $[10^{-3}, 10^3]$.

Impact of the Biot number: The results of the calculations of yield factor γ when Biot number β and Thiele module σ is changing and the porosity $\epsilon_t = 0.1, 0.5, 1$ are depicted in Figure 5.

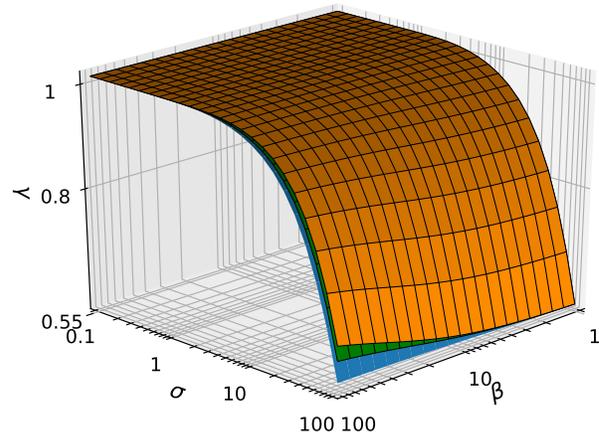


Figure 5. The yield factor γ vs. Biot number β and Thiele module σ at porosity ϵ_t : 0.1, 0.5, 1, the other parameters are as in (40).

In Figure 5, the yield factor γ is presented when concentration \tilde{S}_0 is equal to one. It is apparent that maximal efficiency can be achieved only with Thiele module $\sigma < 1$. It is worth to mention that porosity can be considered as neglected for yield factor. With large values of Thiele module, the yield factor γ is nearing to zero independently of the porosity ϵ_t or β . The yield factor γ , practically, does not depend on β for $\sigma < 1$.

The Biot number and substrate concentration is neglected to yield factor γ with Thiele module $\sigma < 20$. On the other hand, the Nerst diffusion layer may be neglected when the Biot number is higher than approximately 20 [34][37].

Impact of the substrate concentration: Calculation results of the yield factor γ when substrate concentration S_0 and Thiele module σ is changing and the porosity $\epsilon_t = 0.1, 0.5, 1$ are depicted in Figure 6.

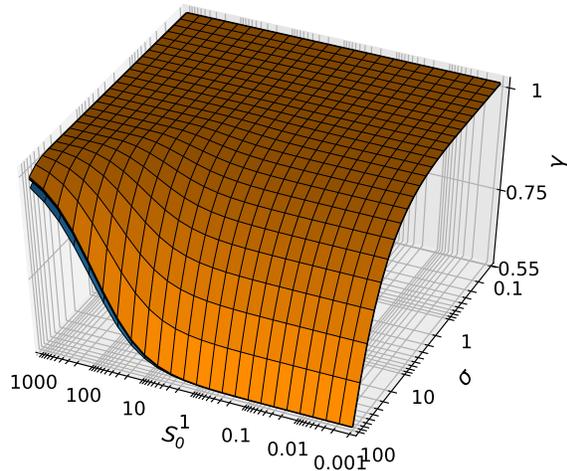


Figure 6. The yield factor γ vs. Thiele module σ and initial concentration S_0 at porosity ϵ_t : 0.1, 0.5, 1, the other parameters are as in (40).

One can see (Figure 6) a non-linear impact of the substrate concentration on the yield factor. As a function of \tilde{S}_0 , the yield factor γ is a monotonously increasing function with the limit of one. At high values of the Thiele module $\sigma > 20$, the yield factor is decreasing to zero near low values of substrate (see Figure 6). The yield factor, limiting value of one, can be reached with the low Thiele module values $\sigma < 20$.

The yield factor γ is, practically, invariant to changes in the substrate concentration \tilde{S}_0 when the Michaelis-Menten kinetics approach the first order ($\tilde{S}_0 \ll 1$) or zero order kinetics ($\tilde{S}_0 \gg 1$). In intermediate values of \tilde{S}_0 , when the kinetics changes from the first to zero order, the yield factor γ noticeably increases with increasing the substrate concentration at low Thiele module values.

Impact of the Thiele Module: It can be seen in Figure 5 and Figure 6 that the yield factor γ , practically, does not depend on σ and approaches to one when the bioreactor acts notably under the bioreaction control ($\sigma < 20$). At mixed conditions when the diffusion action is influenced by both the enzyme kinetics and the diffusion, the yield factor γ noticeably decreases decreasing the substrate concentration S_0 . Figure 6 also shows that the factor γ increases while increasing the substrate concentration (as in Figure 6) as well as when increasing the Biot number β (as in Figure 5).

C. Impact of the porosity

To investigate the dependence of the yield factor γ , the effectiveness factor η on porosity ϵ_t , the calculations were performed in a range of [0.1, 1]. The ranges of Thiele module σ , Biot number β and tortuosity was used as in the previous Section.

Figure 7 shows the effectiveness factor η as an increasing function of the porosity ϵ_t when Thiele module is one and Biot number is two. When the Michaelis-Menten kinetics approaches the first order ($\tilde{S}_0 \ll 1$), the effectiveness factor η becomes a linear function from the porosity. The effectiveness factor η approaches one as increased concentrations are becoming zero order kinetics ($\tilde{S}_0 \gg 1$). Those limits can be

validated from calculated factors at zero order (33) and the first order kinetics (39).

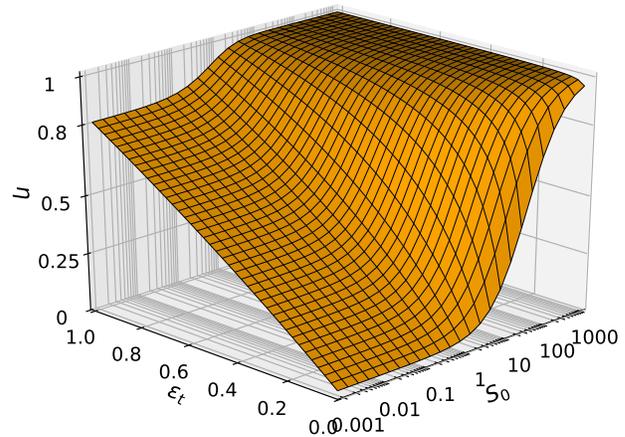


Figure 7. The effectiveness factor η vs. initial concentration S_0 and porosity ϵ_t , the other parameters are as in (40).

Figure 8 shows the effectiveness factor η as an increasing function of the porosity ϵ_t when Thiele module and concentrations are equal to one. The effectiveness factor η , practically, does not depend on the Biot number for $\beta > 2$ and increases with increasing the porosity and the Biot number.

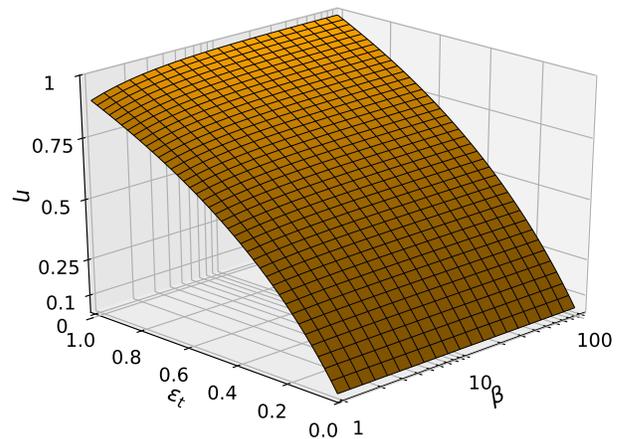


Figure 8. The effectiveness factor η vs. Biot number β and porosity ϵ_t , the other parameters are as in (40).

Figure 9 shows the yield γ as an increasing function of the porosity ϵ_t when concentrations are equal to one and the Biot number is two. It is evident that high Thiele module values $\sigma \gg 10$ are significantly decreasing any efficiency of the system and the yield factor practically does not depend on the porosity ϵ_t . On the other hand, the small Thiele module values give a linear dependency from the porosity. It can be noted that yield as a function of Thiele module is nonlinear and increases with decreasing diffusion module.

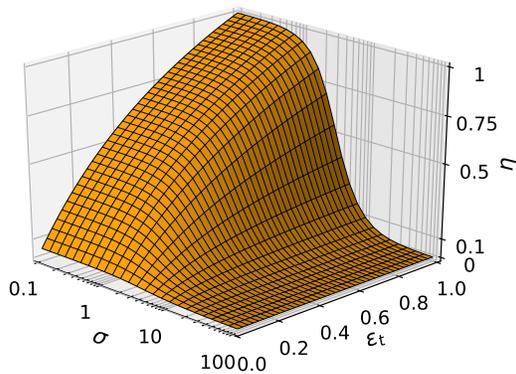


Figure 9. The effectiveness factor η vs. the Thiele module σ and porosity ϵ_t , the other parameters are as in (40).

Figure 10 shows the yield γ as an increasing function of the Thiele module and practically does not depend on the porosity ϵ_t when Thiele module is one and the Biot number is two.

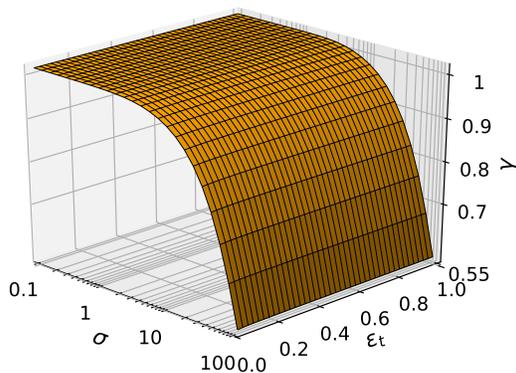


Figure 10. The yield factor γ vs. the Thiele module σ and porosity ϵ_t , the other parameters are as in (40).

VI. CONCLUSION AND FUTURE WORK

The mathematical model (3)-(10) and the corresponding dimensionless model (15)-(23) of the microbioreactor can be successfully used to investigate the behaviour of the catalytic microreactor and to optimize its configuration.

The yield of the product increases with increasing the substrate concentration (Figure 6) and with decreasing the Thiele module (Figure 5). However, the effectiveness of the yield factor can be achieved to the limit of one, with not high Thiele module value $\sigma < 20$. More importantly, the yield factor, practically, does not depend on the Biot number (Figure 5) and the porosity (Figure 10).

The increase in the substrate concentration becomes ineffective when the enzyme reaction is under control of reaction control $\sigma < 20$ (Figure 6). The high yield can be achieved

only when the enzyme kinetics controls the bioreactor action $\sigma < 20$ (Figure 6).

The transient effectiveness factor is an increasing function of porosity and, practically, does not depend on the Biot number (Figure 8). However, strong non-linear dependencies appear when analysing the impact of porosity and Thiele module (Figure 9). Also, is evident that the effectiveness factor approach zero under diffusion control $\sigma > 10$ and is limited of porosity under action control $\sigma < 0.1$. The dependency from substrate concentration (Figure 7) also shows a non-linear relationship. The effectiveness increases to one with increase in the substrate concentration and reduces to the linear function of porosity when $S_0 < 0.01$.

Such formulation can be useful to find the optimal parameters of such biosystem construction [12]. More importantly, it might improve the design and production of microbioreactors.

There are some limitations worth to mention that might be used for the future investigations. First of all, in physical experiments, pellets can not be considered as the perfect spheres, which requires the modelling of more sophisticated domains in 2D and 3D spaces. Secondly, the system with time dependent characteristics should be considered in the future work.

REFERENCES

- [1] L. Petkevicius and R. Baronas, "Numerical Simulation and Analysis of Enzyme-Catalysed Substrate Conversion in a Microbioreactor," in *SIMUL 2017: The Ninth International Conference on Advances in System Simulation*. IARIA, 2017, pp. 1–6.
- [2] A. Regalado-Méndez, R. R. Romero, R. N. Rangel, and S. Skogestad, "Biodiesel Production in Stirred Tank Chemical Reactors: A Numerical Simulation," in *New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering*. Springer, 2015, pp. 109–116.
- [3] O. N. Ağdağ and D. T. Sponza, "Anaerobic/aerobic treatment of municipal landfill leachate in sequential two-stage up-flow anaerobic sludge blanket reactor (UASB)/completely stirred tank reactor (CSTR) systems," *Process Biochemistry*, vol. 40, no. 2, pp. 895–902, 2005.
- [4] A. Miyawaki, S. Taira, and F. Shiraishi, "Performance of continuous stirred-tank reactors connected in series as a photocatalytic reactor system," *Chemical Engineering Journal*, vol. 286, pp. 594–601, 2016.
- [5] M. A. Dareioti and M. Kornaros, "Effect of hydraulic retention time (HRT) on the anaerobic co-digestion of agro-industrial wastes in a two-stage CSTR system," *Bioresour Technol*, vol. 167, pp. 407–415, 2014.
- [6] A. J. Straathof and P. Adlercreutz, *Applied Biocatalysis*. CRC Press, 2000.
- [7] P. M. Doran, *Bioprocess Engineering Principles*. Academic Press, 1995.
- [8] J. Villadsen, J. Nielsen, and G. Liden, *Bioreaction Engineering Principles*. Dordrecht: Springer, 2011.
- [9] S. K. Dubey, A. Pandey, and R. S. Sangwan, *Current Developments in Biotechnology and Bioengineering: Crop Modification, Nutrition, and Food Production*. Elsevier, 2016.
- [10] J. W. Wong, R. D. Tyagi, and A. Pandey, *Current Developments in Biotechnology and Bioengineering: Solid Waste Management*. Elsevier, 2016.
- [11] A. Konti, D. Mamma, D. G. Hatzinikolaou, and D. Kekos, "3-Chloro-1, 2-propanediol biodegradation by Ca-alginate immobilized *Pseudomonas putida* DSM 437 cells applying different processes: mass transfer effects," *Bioprocess and Biosystems Engineering*, vol. 39, no. 10, pp. 1597–1609, 2016.
- [12] D. Caşcaval, A. C. Blaga, and A.-I. Galaction, "Diffusional effects on anaerobic biodegradation of pyridine in a stationary basket bioreactor with immobilized *Bacillus* spp. cells," *Environmental Technology*, pp. 1–13, 2017.

- [13] R. Karande, A. Schmid, and K. Buehler, "Applications of multiphase microreactors for biocatalytic reactions," *Organic Process Research & Development*, vol. 20, no. 2, pp. 361–370, 2016.
- [14] K. F. Jensen, "Flow chemistry—microreaction technology comes of age," *AIChE Journal*, vol. 63, no. 3, pp. 858–869, 2017.
- [15] F. M. Pereira and S. C. Oliveira, "Occurrence of dead core in catalytic particles containing immobilized enzymes: analysis for the michaelis-menten kinetics and assessment of numerical methods," *Bioprocess and Biosystems Engineering*, vol. 39, no. 11, pp. 1717–1727, 2016.
- [16] M. B. Kerby, R. S. Legge, and A. Tripathi, "Measurements of kinetic parameters in a microfluidic reactor," *Analytical Chemistry*, vol. 78, no. 24, pp. 8273–8280, 2006.
- [17] J. Wang, *Analytical Electrochemistry*, 3rd ed. Hoboken, New Jersey: John Wiley & Sons, 2006.
- [18] R. Baronas, F. Ivanauskas, and J. Kulys, *Mathematical Modeling of Biosensors*. Dordrecht: Springer, 2010. [Online]. Available: <http://www.springer.com/chemistry/book/978-90-481-3242-3>
- [19] D. Britz and J. Strutwolf, *Digital Simulation in Electrochemistry*, 4th ed., ser. Monographs in Electrochemistry. Cham: Springer, 2016.
- [20] M. F. Chaplin and C. Bucke, "The large-scale use of enzymes in solution," *Enzyme Technology*, pp. 138–166, 1990.
- [21] H. Gutfreund, *Kinetics for the Life Sciences*. Cambridge: Cambridge University Press, 1995.
- [22] L. A. Segel and M. Slemrod, "The quasi-steady-state assumption: a case study in perturbation," *SIAM Rev.*, vol. 31, no. 3, pp. 446–477, 1989.
- [23] P. Grathwohl, *Diffusion in natural porous media: contaminant transport, sorption/desorption and dissolution kinetics*. Springer Science & Business Media, 2012, vol. 1.
- [24] M. Velkovsky, R. Snider, D. E. Cliffler, and J. P. Wikswa, "Modeling the measurements of cellular fluxes in microreactor devices using thin enzyme electrodes," *J. Math. Chem.*, vol. 49, no. 1, pp. 251–275, 2011.
- [25] L. A. Belfiore, *Transport Phenomena for Chemical Reactor Design*. Hoboken, New Jersey: John Wiley & Sons, 2003.
- [26] D. D. Do and P. F. Greenfield, "The concept of an effectiveness factor for reaction problems involving catalyst deactivation," *Chem. Eng. J.*, vol. 27, no. 2, pp. 99–105, 1983.
- [27] P. M. Doran, *Bioprocess Engineering Principles*, 2nd ed. Waltham, MA: Academic Press, 2013.
- [28] P. Harriott, *Chemical Reactor Design*. New York: Marcel Dekker, 2003.
- [29] H. J. Vos, P. J. Heederik, J. J. M. Potters, and K. C. A. M. Luyben, "Effectiveness factor for spherical biofilm catalysts," *Bioprocess Eng.*, vol. 5, no. 2, pp. 63–72, 1990.
- [30] T. Schulmeister, "Mathematical modelling of the dynamic behaviour of amperometric enzyme electrodes," *Sel. Electrode Rev.*, vol. 12, pp. 203–260, 1990.
- [31] ———, "Mathematical modelling of the dynamic behaviour of amperometric enzyme electrodes," *Selective Electrode Reviews*, vol. 12, no. 2, pp. 203–260, 1990.
- [32] D. J. Fink, T. Na, and J. S. Schultz, "Effectiveness factor calculations for immobilized enzyme catalysts," *Biotechnology and Bioengineering*, vol. 15, no. 5, pp. 879–888, 1973.
- [33] R. Aris, *Mathematical Modeling: a Chemical Engineer's Perspective*. Academic Press, 1999, vol. 1.
- [34] M. E. Davis and R. J. Davis, *Fundamentals of chemical reaction engineering*. New York: McGraw-Hill, 2003.
- [35] D. A. Edwards, B. Goldstein, and D. S. Cohen, "Transport effects on surface-volume biological reactions," *J. Math. Biol.*, vol. 39, no. 6, pp. 533–561, 1999.
- [36] W. T. Vetterling, *Numerical Recipes Example Book (C++): The Art of Scientific Computing*. Cambridge University Press, 2002.
- [37] R. Baronas, "Nonlinear effects of diffusion limitations on the response and sensitivity of amperometric biosensors," *Electrochimica Acta*, vol. 240, pp. 399–407, 2016.

GeSCo: Exploring the Edge Beneath the Cloud in Decentralized Manufacturing

Badarinath Katti*, Christiane Plociennik†, and Michael Schweitzer‡

*Technische Universität Kaiserslautern

Kaiserslautern, Germany

Email: katti@rhrk.uni-kl.de

†DFKI GmbH

Kaiserslautern, Germany

Email: christiane.plociennik@dfki.de

‡SAP SE

Walldorf, Germany

Email: michael.schweitzer@sap.com

Abstract—Decentralized manufacturing is an active research topic in current smart and open integrated factories, and is probably also the future state of practice in both the process and manufacturing industries. The Manufacturing Execution System (MES) is a comprehensive automation software solution that coordinates all the responsibilities of modern production systems. However, the MES solution is essentially designed as a centralized manufacturing control unit, which goes against the principle of the decentralized manufacturing paradigm. When the advantages and downsides of various MES offerings are explored in anticipation of changing production environments, the Cloud MES (CMES) emerges as the most flexible and affordable solution. However, when operated as a cloud based solution, the MES faces another big challenge: connectivity and network latency. To address these problems, we introduce an edge layer called Generic Shop-Floor Connector (GeSCo) near the shop-floor. In other words, the CMES delegates the responsibility of manufacturing control to this edge layer which consequently facilitates decentralization in manufacturing. Finally, the detailed experimental evaluations suggest a marked decrease of the network latency after the introduction of GeSCo layer.

Keywords—Decentralized Manufacturing; Edge Computing; Cloud MES; Cyber Physical Systems; Generic Shop-Floor Connector.

I. INTRODUCTION

A prior version of this work has been published in [1].

Traditionally, the production was conceived to be a layered top-down approach and a corresponding architecture, known as *automation pyramid*, comprising of different layers such as Enterprise Resource Planning (ERP) [2], MES, Supervisory Control And Data Acquisition (SCADA) [3] and shop-floor. This architecture pattern supports the sensors, actuators, numerous software and hardware systems that perform the manufacturing operations on the one hand, and on the other hand also contain management or planning systems that provide access to the enterprise information. There are various flavors of automation pyramid proposed based on different research paradigms. For example, [4] proposes an evolved automation pyramid that supports networked and decentralized production. However, this research work considers the classical automation pyramid (see Figure 1 left) as its base working model.

With the advent of low-cost and smart sensors and subsequently Cyber Physical Systems (CPS), the sensors that are connected to the machines are now *reachable* as they have online capability. Thus, the manufacturing execution systems

can directly co-ordinate with the plant machines. This development has given rise to the possibility of omitting the SCADA layer, the responsibilities of which can be taken over by the manufacturing execution. In many cases, SCADA systems and the connectivity solutions from the MES layer through the SCADA down to the shop floor have been characteristically vendor-specific. They do not follow industry standards and thus make it difficult to replace machines on the shop floor level. The trend of moving towards standardized communication protocols on all layers of the automation pyramid is also fostering this development of circumvention of the SCADA layer as illustrated in Figure 1.

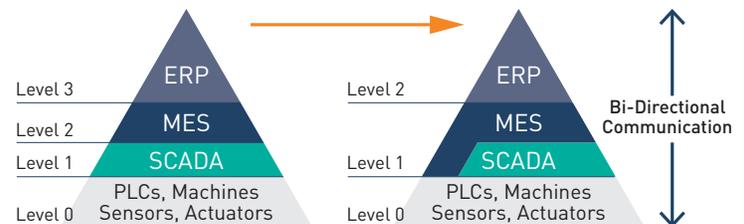


Figure 1: Evolution of classical Automation Pyramid.

In centralized manufacturing, a central entity is responsible for the system planning aimed at the optimization of the objectives of an entire organization [5]. The centralized system is often complex in design and tailor-made to solve a specific class of problems. Centralized systems have slower response times since they employ complex algorithms and analyze more data. In cases of unexpected events and product customizations, centralized systems have proven to be inflexible [5]. Decentralized manufacturing systems are based on distributed control in which the local decision-making bodies react to conditions of the shop-floor at real time. This necessitates the coordination of the supply chain where the operational decisions and activities are shared accurately and in time bound manner with all the entities to avoid uncertainties. However, the solution quality of decentralized systems may be lower since they are based on local information. Furthermore, they require more communication effort.

Both the above-mentioned modes of control in manufac-

turing: centralized and decentralized, are viable depending on the manufacturing circumstances. For instance, the centralized manufacturing is preferred when the manufacturing process involves complex but static procedures [6] such as the determination of the best locations for a set of warehouses and cross-docks. The reasoning behind this choice is that a centralized decision-maker arrives at a decision not based on a certain local maximum, but on a global solution.

When the manufacturing process expects dynamic changes during production, decentralized systems are preferred. In terms of robustness, decentralized systems perform better: The failure of the machines at the lower level of the automation pyramid does not cause the whole system to fail. In a typical centralized system, a failure of central entity can cause the catastrophic failure of the entire system [6]. These arguments support the adoption of decentralized control in manufacturing.

The topic of this paper is cloud MES. The MES historically has been a vendor and industry specific solution and hence, is also called by other names such as Collaborative Production Management (CPM) and Manufacturing Operations Management (MOM) [7]. The IEC 62264-3:2016 standard [8] divides the entire MES activities into four functional areas namely production, maintenance, quality and inventory management. Typical functionalities of production management in MES include sequencing the operations, monitoring the production and to determine the states of different entities involved in production with respect to real time. The focus of this research work is the production management aspect of MES. Although decentralization is the norm of smart manufacturing, MES is conceptually a centralized control system. This does not auger well for current trend of production automation and data exchange methodologies. Therefore, MES should also evolve to adopt the new innovative manufacturing techniques. In the context of manufacturing controlled by centrally managed cloud MES, the research target is to designate the task of production control to the edge layer near the production site. Subsequently, this edge layer enables decentralized decision-making in production control.

This paper explores the various flavors of MES and lists the advantages and the downsides of each of the MES types. It argues that among the various MES, the cloud based MES offers a range of affordable functionality without the problem of vendor lock-in. However, when the MES shifts from on-premise to cloud, it faces the challenge of remote resource management and production control. To enable decentralized production, we propose to introduce an edge layer called *GeSCo* between CMES and shop-floor that coordinates with the shop level entities to perform the task of Production Order (PO) execution. Since modern industries increasingly make decisions by coordinating with business systems, this results in higher network load and latency. To counter these cloud related performance issues, *GeSCo* also caches the routing details and other production related data of CMES.

The outline of the paper is as follows: Section II lists related work. Section III explores various MES adoptions across the industries to give the audience sufficient background before introducing the research question. Section IV describes the use case of the paper and Section V highlights the problem of network latency in the context of high speed manufacturing. Section VI introduces an edge layer that acts as production control delegate. It then also describes the responsibilities such

an edge layer should fulfil in Section VII. Section VIII presents a system architecture that addresses these challenges. Section IX presents implementation details based on the proposed system design and some simulation results. Section X provides the conclusion and an outlook on the future work.

II. RELATED WORK

This section is divided into two parts. The subsection II-1 catalogues the previous and contemporary research in the area of decentralized manufacturing, where as the subsection II-2 details the specialized work in the domain of cloud manufacturing.

1) *Edge Analytics and Decentralized Manufacturing*: Edge computing is in practice since two decades and is also known by other names such as fog computing, mobile edge computing, cloudlets and cyber foraging [9]. Edge analytics applied to the domain of manufacturing addresses the problem of network latency and enables to take decisions at runtime in production and thus, can adopt to changes in the PO within short time. Moreover, [10] declares that owing to the rigidity and low receptiveness to changes in the manufacturing, centralized manufacturing practices were replaced by decentralized manufacturing models. [11] proposes decentralized work-in-progress manufacturing control that serves as an alternative to the centralized manufacturing systems. The RFID-enabled MES was introduced for mass-customization in manufacturing that faced challenges of manual and paper-based data collection, production plans and schedules [12]. However, the assumption was that machines in the factory shop-floor are at best partially connected and the decision-making rests entirely on employees on the shop-floor. Agent-based manufacturing [13] and holonic manufacturing [14] introduced the concept of artificial intelligence in manufacturing with an aim to respond promptly and correctly to changes in PO. [15] professes the idea of edge datacenters that process the data on behalf of IoT devices and delegate to the cloud only when more complex analysis is required. [16] recognizes the issues relating to cloud computing such as latency and low Quality of Service (QoS), and argues that edge computing is the solution. It also proposes an extensible edge server architecture as an ongoing work. [17] proposes a Centralized Scheduling System (CSS) and decentralized MES, where the latter follows a fixed global schedule and turns to CSS in case of perturbation. [18] discusses the autonomous MES that generates alternative schedules when given schedule is infeasible. However, [19] argues that localization of decision-making with an obligation to decentralize has the risk of losing the global vision of the network. [20][21] argue that even though the decentralization of manufacturing is the norm in the future, there are cases where a centralized entity is obligatory to overwrite the lower level decisions, e.g., in the event of redefinition of production processes at higher levels of automation pyramid. [22] also contends that the absence of a central decision-making body necessitates continuous harmonization of objectives among the agents leading to high coordinative complexity. Therefore, there is a renewed interest in incorporating centralized production control concepts to manufacturing.

2) *Cloud Manufacturing*: There have been several works, for example [23][24], in the domain of cloud manufacturing, that combine the emerging advanced technologies, such as cloud computing, virtualization, internet of things and service

oriented architecture. In a broad categorization, two types of cloud computing adoptions in the manufacturing are proposed, namely, direct adoption of cloud computing in manufacturing and centralized management of distributed resources that are encapsulated as cloud services [25]. The latter categorization is also known as distributed manufacturing. The potentials and relationships among cloud computing, internet of things and cloud manufacturing is investigated in [26]. [27][28] illustrate the concept of centrally managed CMES, but its application area is distributed manufacturing, which is outside the purview of this paper. [15] also argues that even though cloud data-centers provide cheaper and unlimited computing power, the fundamental practice of storing the manufacturing control and associated exception handling data necessary to successfully carry out production in the cloud datacenter is increasingly being challenged due to rapidly growing requirement of making production decisions with minimum data processing delays and data transfer to facilitate smart manufacturing. There is no research work that focuses on enabling the edge computing when CMES is in control of production to counter the problem of connectivity and network latency. In general, the research focus in the domain of manufacturing has shifted from centralized manufacturing systems - and MES in particular - to the decentralized paradigm of manufacturing. This research paper is novel in the aspect that it focuses on the adaptation of CMES, which is traditionally linked to the centralized paradigm, to the context of decentralized manufacturing. In other words, it attempts to retain a degree of centralized aspects of manufacturing to strike the right balance.

III. MES TAXONOMY

This section explores the various available generic MES offerings that are adopted across the manufacturing industries. In this industry-neutral MES study, the advantages and the downsides of each of the MES type are weighed up in order to make a applicability assessment from a manufacturer point of view.

1) *In-House MES solution*: This approach involves implementation of a customized MES that fits to the specific needs of the manufacturer. It involves direct interaction of the users and developers. The manufacturer should own a group of business analysts and developers with a common reporting line to facilitate smooth coordination between the teams. These human resources should have vast experience in IT implementation and a thorough understanding of the business processes. The home-grown MES allows the user to have an in-depth knowledge of the system functionality and hence, complete control over the manufacturing processes. The organization has proprietary rights of the software and also possesses the knowledge that was gained during software development. The responsibility of software maintenance over a long period of time rests with the manufacturer. At the same time, the organization should adapt to changing business requirements and newer technologies. A MES is inherently difficult to own and maintain and even more rigid to evolve owing to the tight coupling of IT infrastructure to the manufacturing operations [12]. This characteristic aversion of MES to change quickly also hinders the implementation of a streamlined production process. The inability of early adoption of new innovative technologies adversely impacts the revenue generation of the manufacturer.

2) *Proprietary production control system*: The proprietary production control system that is part of the automation hardware is another facet of MES. These production control systems convert the ERP orders to technical production orders for the assembly lines. This allows the manufacturer to do away with the development of the software and hence, lessen the cost burden. It also enables the manufacturer to immediately focus on the production. However, such production control systems are tightly coupled to the machinery and hence, even a small change in production creates a ripple effect across the automation layers. When the hardware and subsequently the production control software is discontinued, future manufacturing maintenance is not safeguarded. This compromises the flexibility and future security of the entire plant.

3) *Third Party Vendor MES*: A third classification is the third-party vendor MES which is built according to the functional model specified by the manufacturer. The vendor guarantees long term maintenance and further development of MES modules, and integrate future customer requirements in the product design and development. The selection of MES generally results in long term relationship with the MES vendor in the interest of protection of investment.

The vendor provides a proven and off-the-shelf solution that incorporates industry best practices, along with professional support and training to the work force of the manufacturing organization. On the other hand, the continuation of the status-quo after successful installation of MES is expensive since it involves upgradation of hardware components and IT solutions owing to their short innovation cycles. The additional difficulties such as platform dependency, license model and work force that needs to be trained to use the software come to the fore and further increase the cost pressure on the manufacturers. Moreover, these custom built MES command a high price which is difficult to justify for some of the manufacturers. To that end, a detailed analysis of investment is necessary taking into account the life cycle and cost of maintenance during the feasibility evaluation of an MES vendor.

4) *Cloud based MES*: To address the above described difficulties, the traditional MES should be replaced by a comprehensive MES setup that can quickly adapt to newer innovative technologies and offer significant cost benefits to the manufacturer at the same time. The cloud based MES [24] is one such solution. The cloud based MES is a blend of various IT technologies such as distributed computing, internet technology, hardware virtualization and open source software. To be more precise, Internet of things (IoT), which enables the perception, internet connection, acquisition and automatic control of various manufacturing resources and capabilities, is the core enabling technology for the implementation of cloud based MES [24]. Cloud based solutions, in general, are best described as web based solutions that run on remote servers and accessed over the internet via standard web browsers [2]. Cloud MES solutions are offered as IaaS (Infrastructure as a service), PaaS (Platform as a service) and SaaS (Software as a service) layers in the cloud architecture that are demand driven and charged as per usage [29].

The services in cloud based MES are generated by virtualizing and encapsulating the perceived manufacturing resources and capabilities [30]. These MES solutions are mostly assembled from configurable software components. The generic set

of functionalities is built as per the customers requirements and typically, the functionalities provided by cloud based MES are richer than on-premise counterparts [26] and are also simple, fast and cheap [31]. Another main benefit of the cloud based MES is that it requires nearly no IT resource investment [2]. This lowers the entry costs for smaller firms that try to benefit from compute-intensive business analytics that were previously available only to large corporations. This also lowers the IT barriers to innovation in the manufacturing processes [26]. The cloud based MES helps smoothly face peak production demand without additional investment on on-premise IT resources [32]. This is made possible with the virtualization principle of cloud computing technology. The cloud virtualization facilitates multiplexing of a physical equipment by a privileged hypervisor kernel, thereby providing the end-users separate environments to execute their applications. The argument of virtualization holds true also in cases of redundancy or upgrade costs of the on-premise resources. The dearth of skilled resources that are acquainted with MES technology, achieving the ROI and technology compatibility are no longer the problems in the cloud scenario. Since the cloud servers are run as per the necessity, licenses can be increased or decreased accordingly. This decision need not be made upfront.

IV. USE CASE

The communication between the both on-premise and cloud MES can roughly be described as follows. During production execution, the shop-floor constantly seeks information from MES. The work stations at the shop-floor request MES for routing details at every stage of the production. Each work station collects the operation, Bill Of Materials (BOM), machine parameters and other resource configuration details. Once this information is collected the machine is instructed on how to proceed with that step of the production process. Once that step of the production is completed, the work station informs MES the same along with the generated results. The MES then processes the results and accordingly sets the next operation of the production. This process continues until all the planned operations are executed to manufacture the planned component. During exceptional cases or conflicting goals, if the need arises, the routing path is changed, as instructed by MES, to accommodate the exceptional situations. For example, the work in progress is diverted to rework station if the concerns regarding the quality of the products are raised.

V. CHALLENGES IN CLOUD BASED MES

Section III-4 has presented the cloud based MES and its advantages over the classical MES. Nevertheless, there are certain challenges in the cloud MES, or cloud computing technology in general. The cloud downtime and network latency are critical concerns for the manufacturer. This latency becomes even more challenging in high speed manufacturing scenarios where the right information is required at the right time. It is a difficult proposition to measure the exact latency of the network. Historically, it depends on the number of router hops between the client application and the target machine. The network latency is fairly measurable or predictable when the intervening routers are governed by the same corporation. The situation changes when the business migrates to cloud and the issue of latency becomes increasingly complex. The notion of enterprise data centers is no longer followed. The

nature of applications is changing from being contained in a local infrastructure within an organization to distributed across the world. Since these applications are deployed across the world, they have varying degrees of latency that are based on each of the internet connections. Hence, the location of data centers plays a significant role in determining the network latency. Furthermore, the network latency is a function of internet traffic that undergoes random fluctuation for the same bandwidth and infrastructure.

The loss of governance is perceived as another biggest impediment to acceptance of cloud based manufacturing solutions [33]. When business applications are moved to cloud, it forces the organizations to accept the control of the service provider on several important issues and areas of business and manufacturing data. As a result, the cloud solution provider will have overarching influence on the business processes. The fact that the valuable enterprise data resides outside the company firewall raises serious security and privacy concerns. However, the security concerns of the cloud based MES are outside the purview of this research work.

The communication between traditional MES and shop-floor takes place over WAN, which means that the transmission delay is not bounded [34]. When moving from MES to CMES, network latency becomes an even bigger challenge as the geographical distance and, consequently, the number of intermittent routers increase. The request and response data travel through the source and destination entities in the network via a series of routers. These data packets suffer several types of delays at each node along the network path. A node can be a source, destination or an intermittent router. The following are significant delays encountered by the data packets:

- Nodal processing delay
- Queueing delay
- Transmission delay
- Propagation delay

The throughput of the network is greatly affected by these network delays, particularly the nodal queueing delay and the propagation delay [35]. The delays are explained in the context of Figure 2. The data packets are sent from source to destination via routers r_1 and r_2 . Each router has an incoming queue and an outbound link to each of the connected routers. The packet arriving at a router goes through the queue and the router determines the outbound link after examination of the packet header. An incoming data packet is immediately bound to outbound link if the router queue is empty and there are no packets being sent on the outbound link at the time. If the router queue is non-empty or the corresponding outbound link is busy, the incoming packet joins the router queue. When the data packet arrives at a router, the router examines the packet header for redirection to the appropriate destination. This causes a delay which is known as Processing delay d_{proc} and is the key component of network delay. The node also checks for bit level errors in the packet arising while transmitting from the previous node. After this nodal processing, the router directs the packet to a queue that precedes the outbound link. Normally, the processing delay is of the order of few tens of microseconds in most of the high processing routers in case of forwarding a simple packet [35]. But this delay can be up to a millisecond in case the router undertakes the task of performing encryption algorithms aimed at examining or

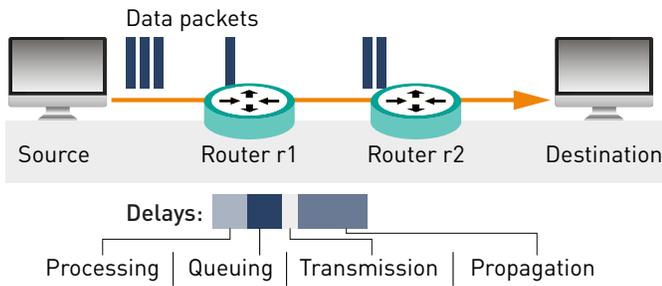
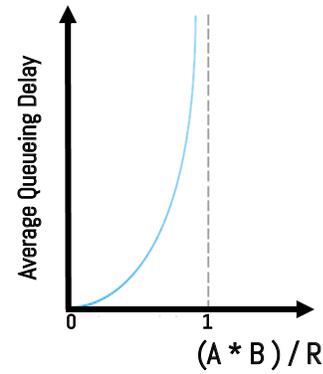


Figure 2: Illustration of network delays.

modifying a packet designed for security and legal aspects [35]. The time a packet spends in the queue while earlier packets are transmitted at the node is called queueing delay d_{queue} . The incoming packet experiences zero queueing delay when the router queue is empty and no other packet is being transmitted by the router. Alternatively, the incoming packet experiences a queueing delay in direct accordance with the length of the router queue. The router transmits the data at a rate known as transmission rate. When the data packets arrive for a sustained period at a given router at a rate more than its transmission rate, these data packets will queue in at the router.

To gain some insight, let A denote the average number of packets that arrive at the router queue per unit time. Let R be the transmission rate of the router; that is, it is the rate at which the bits are pushed out of router queue. For the sake of simplicity, suppose all the packets consist of B number of bits. Then, the average number of bits that arrive per unit time at the router queue is $(A * B)$. The ratio of $(A * B)/R$, called network traffic intensity, plays an important role in determining the queueing delay. If network traffic intensity is less than 1, the nature of arriving data packets influences the queueing delay. If a data packet arrives every A/R units of time, each of these packets then arrives at an empty queue and will not encounter the queueing delay. Conversely, if the packets arrive in bursts due to traffic congestion, it then results in substantial average queueing delay. For example, assume P packets arrive simultaneously every $(A/R) * P$ units of time. The first packet that is transmitted will encounter no queueing delay. Nonetheless, second packet encounters a queueing delay of (A/R) units of time. Similarly, the third packet experiences a queueing delay of $2 * (A/R)$. In general, the n^{th} data packet will experience a queueing delay of $(n - 1) * (A/R)$ units of time. However, the packet queueing does not follow a pattern in practical situations and the packets are spaced apart by an arbitrary amount of time. Therefore, the above quantity alone is not adequate to fully characterize the queueing delay. Nevertheless, it is a useful tool in the estimation of queueing delay. If the traffic intensity tends to 0, approximation is then few packets arrive and they are spaced far apart in time. The probability that each packet encounters a non-empty queue is close to 0. Conversely, when the traffic intensity tends to 1, there will be intervals of time when the packet arrival rate is greater than the network transmission capacity. The packet queue at the router site builds and grows open-endedly when the packet arrival rate is greater than the router transmission rate. On the contrary, when the packet arrival rate is less than

Figure 3: Dependence of average queueing delay d_{queue} on traffic intensity [36].

the router transmission rate, the size of the queue shrinks.

The qualitative dependence of average queueing delay on the network traffic intensity is demonstrated in Figure 3. It can be observed from Figure 3 that as the traffic intensity tends to 1, the average queueing delay grows exponentially. When the packet arrival rate is greater than router transmission rate, the size of packet queue grows at the router. However, this cannot continue indefinitely due to the finite capacity of the router queue. Therefore, the router drops the packet when it finds no place at its queue. Such a dropped packet is lost and this phenomenon is called *Packet Loss*. At this juncture, the client that transmitted the packet to the network core expecting the delivery acknowledgement from the server re-transmits the packet after waiting for a specified amount of time. This reduces the throughput of the network connection. The ratio of lost packets depends on the router queue capacity, network traffic intensity and the nature of traffic arriving at the queue. In general, the queue capacity greatly depends on the router design and cost. From the above discussion, it can be inferred that the network latency not only depends on the delay during transmission, but also on the packet loss.

The router takes a finite time to transfer the bits of a data packet onto the outbound link. This time is known as transmission delay d_{trans} and mathematically, it is defined as B/R . It is directly proportional to the number of bits in the data packet and independent of queue length and distance between the two nodes. It is of the order of microseconds [35].

The packet on the outbound link propagates to the next node in a time known as the propagation delay. If l is the length of the physical link and v is the propagation speed of the data packet in the physical link, the propagation delay d_{prop} is then given by l/v . The propagation delay varies directly with the distance between the adjacent nodes and is of the order of tens of microseconds to milliseconds [35]. It can vary significantly between few microseconds for a link connecting two routers within the same intranet to millisecond for a link joining two routers thousands of kilometers apart.

The total nodal delay d_{nodal} is then given by [35]

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop} \quad (1)$$

If there are N number of similar routers between the source and destination spaced apart at equal distances, then the end-

to-end delay $d_{end-to-end}$ is measured as

$$d_{end-to-end} = N * (d_{proc} + d_{trans} + d_{prop}) + \sum_{n=1}^N d_{queue_n} \quad (2)$$

where the last part of the above equation is sum of the queuing delays experienced at each of the routers. The network delays are directly proportional to the distance and consequently, the number of intermittent routers, between the client and the server. In practice, with the exception of d_{proc} , which is on the order of microseconds, all other above-mentioned delays are on the order of milliseconds [36].

It is not possible to accurately determine the latency between two fixed points since the data packets encapsulated at the network layer of OSI model need to pass through several proprietary routers of the internet before reaching the destination. Each of these routers has unpredictable traffic, which is dependent on variety of factors and hence, the network latency is a function of internet traffic that undergoes random fluctuation for the same bandwidth and infrastructure. Therefore, instead of imposing hard real-time constraints, the practical unit of measurement should be average time for the network latency.

The virtualization principle of cloud computing that can be applied at different levels such as computer hardware, operating system, storage and network also introduces its own series of packet delays and causes further performance degradation.

Figure 4 illustrates this situation where there are three operations - welding, color spraying and quality check, which are required to be performed to produce the planned component. In the state of the art industries, the work stations constantly communicate with CMES to seek process parameters, recipe, machine configuration values and push the results during production control. The problem of network latency which is encountered each time the request is created to fetch the next operation details from CMES does not auger well in high speed manufacturing scenarios.

In addition, although cloud providers claim near 100% availability, there are instances in the life cycle of cloud solutions where the services are disrupted due to many reasons such as electric failure, hardware failure, cascading failure on routers and cloud downtime arising out of data center migration, server update against vulnerability et cetera. These incidences, on an average, reduce the availability to 99.91%, which in other words a non-availability of 7.884 hours per year [37]. Such network outages are not acceptable in the event of manufacturing a priority order.

VI. INTRODUCING AN EDGE LAYER

To realize the decentralized production, this research paper proposes introducing an edge layer called *Generic Shop-Floor Connector (GeSCo)* between CMES and shop-floor. However, as explained in Section V, the network latency is directly proportional to the geographic distance. The MES in cloud is not guaranteed to be close to the site of production. With this view, the production control data of CMES is cached in proximity to the shop-floor can reduce the problem of network latency. To that end, GeSCo is an ideal place to store the cached data.

GeSCos are close to, but not tightly coupled to the shop-floor. They control the production processes and collect the data to and from the shop-floor and enterprise software. GeSCos also help in enabling the *plug and produce* feature of today's smart factory, since they can connect to wide variety of industry specific data sources of diverse manufacturers, such as OPC UA, classical OPC and http based web services. Due to the physical proximity of GeSCos and shop-floor, the data communication latency is short as data packets need not cross multiple routers. GeSCos also alleviate the problem of latency introduced by the virtualization layer of cloud infrastructure explained in Section V. The cached data constitutes production control data of part of/complete/multiple PO(s). Such information empowers GeSCo to take decisions with regard to production control without the consulting the centralized CMES and hence, it facilitates the implementation of *decentralization* of the production execution.

In its basic conception, the GeSCo is a web service and other numerous industrial communication protocols framework. It collaborates with enterprise software and diverse industrial data sources to execute a PO by performing division of labor in the shop-floor under the supervision of CMES, i.e., it distributes the production operations to resources on the shop floor based on the production recipe at run-time. The introduction of GeSCo in the shop-floor is not to take over the role of SCADA. Instead, it should just serve as a thin client to CMES server. Based on these arguments, the CMES and the shop-floor communication evolution can be illustrated as in Figure 5.

VII. REQUIREMENTS FOR GESCO

Subsequent to the caching of the production control data, the intention is to reduce the communication between the GeSCo and CMES as far as possible. Several exceptional situations may arise in the shop-floor while the GeSCo is in control of the production execution. The manufacturing resource breakdown is one such case in point which is a highly disruptive occurrence in an automated production environment. Even as preventive maintenance or repair is a preferable way to increase the system reliability and significant system cost reduction, [38] claims that, in real-life manufacturing systems the machine breakdowns are inevitable. The GeSCo should anticipate such an eventuality and must be well equipped to take appropriate course of action.

The current manufacturing operation cannot be swapped to another manufacturing resource when there are no alternative manufacturing resources in the shop floor cell. In such a case, the GeSCo should preempt all the other steps of the routing and retain its state. Under such an abort/resume policy in case of random manufacturing resource breakdown, production should resume with the processing of the preempted step of the routing after the breakdown is fixed. When GeSCo has started execution of another PO of different product variant with no dependency on the resource which has broken down, it should resume the execution of aborted PO after completion of the current PO.

In a job shop environment, the presence of multi purpose manufacturing resources enables to execute multiple operations on several alternative resources. In such a scenario, GeSCo must reschedule the production routing by replacing the disrupted resource with an alternative resource. In the event of

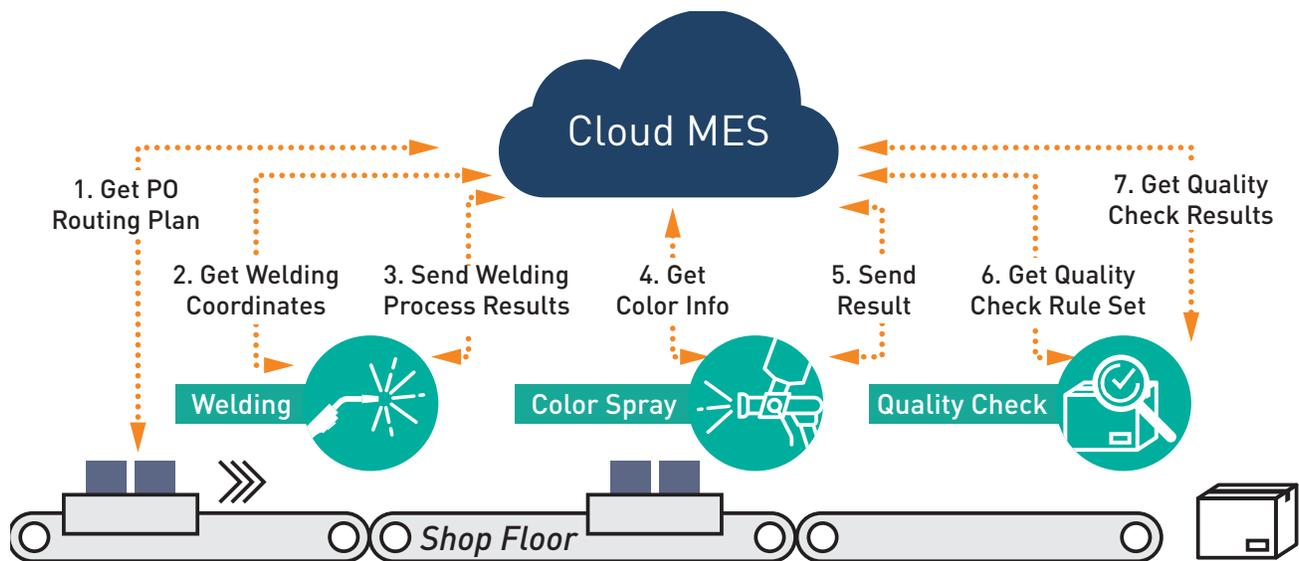


Figure 4: CMES - Shop floor connectivity in production.



Figure 5: Evolution of CMES - Shop Floor Connectivity.

manufacturing resource replacement, the new manufacturing resource should be introduced easily and quickly into the manufacturing system without reprogramming or reconfiguration of the production setup. The GeSCo should be resilient to such *plug and produce* concepts.

When the PO is changed at ERP during the execution, the CMES should deliver the necessary information promptly to GeSCo. The GeSCo should check the feasibility of the changed PO, take appropriate measures and convey the same to the upper layers of automation pyramid. Under normal circumstances, the GeSCo should adopt the First-In-First-Out (FIFO) policy for the execution of a PO. However, when the production routing consists of manufacturing operations of different lead times, the priority order in the pipeline should finish the execution at the earliest. Therefore, the provision should be made in GeSCo by defining a priority policy to put a non-priority order on *hold* state in order to expedite the execution of the priority order.

The traditional MES creates static production routing where manufacturing resources and operations are coupled together and pushed down to shop-floor execution. This approach does not allow the edge component the freedom to make decisions at the shop-floor. In case of deviation from the production planning, the edge component seeks the directions

from CMES to recover from the path of deviation to successfully perform all the activities of production. However, in order to provide more autonomy to the GeSCo, the CMES should only create the abstract production planning without tying the manufacturing operations to resources. This process should be performed in GeSCo. The GeSCo should possess local intelligence during the dispatch of manufacturing operations to manufacturing resources. In addition to their *reachable* property, the modern manufacturing resources also known as Cyber Physical Production Systems (CPPS), have more computing power to complement large number of embedded sensors and actuators. These resources can track their state, PO buffer, and are *aware* of their various configurations to manufacture products with unique characteristics. In state-of-the-art factories, the shop-floor is considered to be a *service market place* where different manufacturing requirements are matched against the corresponding services offered by the resources to produce a tailored product defined by the customer.

It should be the responsibility of GeSCo to associate each operation of a PO to a particular manufacturing resource also called CPPS in order to process a semi-finished assembly also called CPS. The dispatched operation characterizes the logical binding between the CPS and CPPS. The changes brought about by this combination of the CPS, CPPS and GeSCo that has the relevant contextual information of the current POs drive changes in manufacturing production and control, and actuate the remodeling of centralized to truly decentralized production decision-making systems.

The idea is not to store complete informational and operational technology information in the GeSCo to make these runtime decisions. The provision should be made where the manufacturing resources publish their capabilities to the GeSCo. The GeSCo should utilize this information to assign a routing step to one of the manufacturing resources. The PO should also push the required abstract services in case of quality non-conformance along with the non-conformance codes. In the event of quality non-conformance, the GeSCo only looks-up the non-conformance code and seeks the corresponding

services from the manufacturing resources. Another major challenge is to make feasible decisions taking into the account the physical configuration of the cells of the shop-floor.

In all the above described exceptional situations, the GeSCo should either resolve or find an alternative course of actions. The objective of this exercise is the successful completion of the production execution. The CMES should support this goal by sending meaningful data at the right time.

A. Challenges of Integration of GeSCo: A Survey

The GeSCo should assume the role of the CMES after the PO is transferred to its cache. The transfer of production control to the GeSCo is smooth under normal circumstances when the production encounters no problems. However, the system should be designed such that it should be robust against production fluctuations and should mitigate or solve the problems that may arise under exceptional circumstances.

In order to determine which responsibilities such a system must fulfill, several experts in the field of manufacturing were asked to prioritize the challenges for GeSCo during the execution of shop orders. The results of this survey are, in descending order of their weighted average:

- 1) Determination of next routing step since business rules that govern the routing decisions are present in the CMES
- 2) Semantic translation of data arriving from CMES to technology and business agnostic solution such as GeSCo
- 3) Adaptation in GeSCo in the event of change of the data model in centralized CMES
- 4) Determination of the suitable resources to perform the current operation
- 5) Routing-path substitution in the event of machine breakdown [13]
- 6) Dealing with the change of the PO [13]
- 7) Handling the POs of high priority [13]
- 8) Course of action in the event of quality defects
- 9) Resumption of production after a disruption due to unforeseen circumstances
- 10) Course of action in the event of unavailability of raw materials
- 11) Distributed manufacturing where components are being manufactured at different sites

VIII. PROPOSED SYSTEM ARCHITECTURE

The solution architecture should be designed taking into account the challenges mentioned in Section VII-A. It should enable the CMES to exercise control over the production process while at the same time ensuring a smooth integration of the GeSCo for providing flexibility in exceptional cases. Hence, the architecture should incorporate both centralized and decentralized aspects.

A. Design of CMES

This section describes the proposed set of building blocks and services that are required in the CMES. The overall architecture is depicted in Figure 7.

1) *Production Planning System*: This application layer enables the human production planner to plan the production sequence in a generic way. To this end, it has different maintenance user interfaces that help define the plant and product definition, operation planning and production execution



Figure 6: Resource Virtualization.

aspects. This master data facilitates the design of BOM and the shop-floor routing for a product variant. This unit also enables the human to create and release the PO to the shop-floor.

2) *Manufacturing Resource Model and Servitization*: Remote resource sharing and management is a challenge to CMES since it is geographically separated from the shop-floor. The resource virtualization is the key idea behind building the cloud services in the context of manufacturing. The resource model is the transformation of a real manufacturing resource to a virtual or logical resource. Each manufacturing resource is modeled formally with a set of inputs and outputs according to its main functionality. The functional and non-functional capabilities of the resource can be semantically modeled. The model is then subjected to real-to-virtual mapping methods to map to a logical resource as illustrated in Figure 6. The concept of enriching the digital plant models by making the virtual copies of the manufacturing resources with near real time data from sensors also makes the information flow more transparent. The virtual resource servitization is the transformation of abstract concepts of capabilities provided by these resources into formal services that are understandable by the cloud platform. This process involves several aspects such as definition of the service model, message model, ports and protocols. The service model includes the template for the service offered by cloud platform. The reception of inputs and generation of outputs of the service is defined in the message modeling process. The port modeling involves the definition of functional operation port used to accomplish the operation target. The protocol binding specifies the different protocols that are supported by the service.

This service interface of virtual resource enables GeSCo to store the resource relevant data in a realistic resource model, also called as resource digital twin [39]. The GeSCo collects the machine data from resource periodically and pushes it to CMES resource model. This assists in real time monitoring of the manufacturing resource for the purpose of tracking the status and understanding its behavior in interaction with other manufacturing systems, and also to calculate the equipment effectiveness. Further, the data is archived and the aggregated historical data is fed to the predictive analytics tool to find the insights into the resource behavior.

3) *Dispatcher*: The PO created and released by the production planner is transferred from the CMES to the shop-floor by the dispatcher. The logic of transferring the priority order(s) is pre-loaded into the dispatcher. The parameters that expedite the release and subsequent transfer to the shop-floor are production end date, priority customer, and inventory and manufacturing resource availability. The GeSCo, introduced in this paper, is a technology and business agnostic solution. Therefore, the dispatcher should send the unambiguous data, for example, a collaborative product definition and operations semantic model to the GeSCo. The GeSCo translates this information to its

compatible data model for further processing.

4) *Data mining and predictive analytics*: Instead of relying on human expertise alone, there is an increasing inclination towards aggregating and processing a large amount of data at the shop-floor, which in turn enables to train better models for classification, clustering and prediction. This component analyzes the current and past semi-structured or unstructured data and extracts useful patterns and transfers this knowledge to GeSCo. This knowledge of past experience is then helpful for GeSCo to take run-time decisions that solve or mitigate the problems arising in the shop-floor during production. This information is also helpful to achieve optimization of the production processes in the shop-floor.

5) *Information systems*: This constituent stores the product genealogy including complete work instructions, components and phantom assemblies, operation flow and routing, manufacturing resources and work centers employed, bill of materials, activities on the shop-floor, rework instructions and the discrepancies. This is realized using the Digital Object Memory (DOMe) [40], which maintains all the information about a product instance over its production lifecycle, where each product is identified and tracked using RFID tag that contains the unique shop-floor control number. Since DOMe is centrally accessible to all the involved entities of production, it enables production coordination among these entities, compilation of historic manufacturing reports, quality investigations and process improvements.

B. Design of GeSCo

The GeSCo should consist of the following components with dedicated responsibilities (see also Figure 7):

1) *Manufacturing Resource Perception Layer*: To achieve harmonization among various manufacturing resources, they need to be coupled together. The perception layer undertakes this responsibility of loose coupling of different resources on the shop-floor. The different manufacturing resources at the site also register themselves to this layer. The registration can take place either with the resource meta-data or the resource endpoint that permits the perception layer to browse the resource data structures to extract the meta-data of the resource. To this end, this module has internally a sub-module known as Capability Discovery Repository (CDR), which stores the capabilities of the various manufacturing resources. The manufacturing resources are also allowed to directly announce all their capabilities semantically to the CDR. However, a formal explicit specification of shared concepts [41] and relationship among those concepts, also called ontology, needs to be modeled at the organization level in order to realize the semantic publishing of the capabilities. A static service which provides access to the created ontologies enables referencing and dereferencing of the semantic concepts. The decentralization facilitator exploits this semantic information from CDR to arrive at the decisions at run-time. To this end, the authors extended OWL-S and SAWSDL specifications to the OPC-UA application specific methods in order to automate the process of method discovery and subsequent method composition in [42] [43].

The perception layer should support the standard industrial communication protocols, such as OPC UA, classic OPC and HTTP based data sources. These IoT protocols are employed to perceive different manufacturing resources with an intent

to enable intelligent identification, detection, communication, tracking, monitoring and management. The effectiveness of this exercise hinges on the ability of this layer to extract the key information from the real resources.

2) *Production Control Data Cache*: This component stores the data delivered by the CMES. It contains the blueprint of the production execution on the shop-floor, which is the detailed routing information in the case of discrete manufacturing. Various entities of GeSCo such as decentralization facilitator and production engine base their decisions and actions on this cached production execution data. This unit is designed to address the first three challenges listed in Section VII-A.

3) *Decentralization Facilitator*: This entity enables the decentralization in manufacturing by coordinating with various manufacturing resources and CMES, and thus helps address the challenge of determining the suitable resources for a particular operation. The layer maintains the virtual resource pool consisting of a collection of virtual manufacturing resources. It is used in run-time classification of resources that aids in on-demand resource capability matching. The virtual resource management helps GeSCo identify capabilities intelligently by semantically searching for suitable services and the manufacturing resources on the shop-floor to meet the production requirement.

4) *Exception Handler*: This block of the GeSCo is accountable for overcoming any shortcomings that arise in the production environment. These shortcomings are explained in Section VII-A, numbers 5 to 9. The exception handler either attempts to find an alternate course of action by local coordination or seeks further instructions from the centralized entity which has global picture of the system.

5) *Production Engine and Work-In-Progress Monitor*: The production engine is the heart of the GeSCo that collaborates with all the other components of GeSCo to achieve the end goal of successful completion of the PO. It fetches the PO information and routing details from the production control data cache and delegates the responsibility of matching the manufacturing resources for the given operation to the decentralization facilitator. After the decision-making process, the production engine delegates the job to the perception layer that assigns the operation to the real resources after the necessary configuration. The production engine also assigns the unique PO identifier to the smart product or the product carrier at the start of the PO, so that the carrier can be identified and tracked any time during production. During the dispatch of each routing step of a PO, the manufacturing operation harnesses the unique CPS identifier and binds the product to the manufacturing resource. The PO is put on hold in the event of non-availability of default and alternate resources, and is only resumed after the required resource registers to the perception layer. To ensure the production is running as expected, it is necessary to monitor run-time status and respond to changes. In case of changes and exceptions, this layer coordinates with decentralization facilitator and exception handler to solve or mitigate the contingency. The production engine also has the intelligence to recognize the situations where GeSCo cannot take the optimal decision based on local information. In such scenarios, it seeks the master data, the singular source of truth, stored in centralized CMES.

6) *Production Process Logger*: This component uploads the variety of knowledge it gathers during the production onto

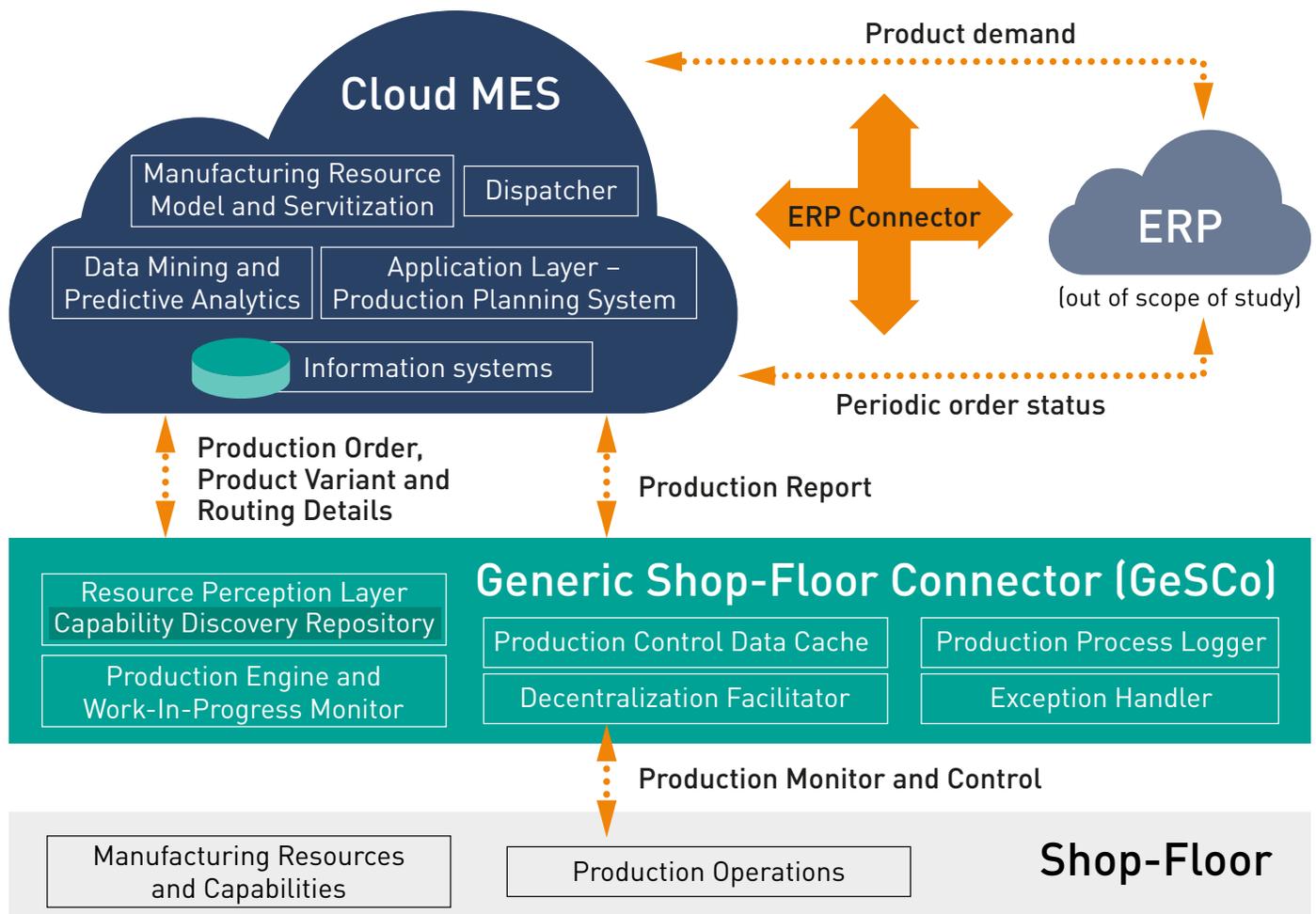


Figure 7: Integration of GeSCo with CMES.

the CMES. This unstructured data is subjected to analysis and an effort is made by CMES to find patterns and transform it into a structured data. This knowledge in turn can be channeled as a feedback to the closed loop system in order to optimize the production in the long run.

IX. IMPLEMENTATION

In order to demonstrate the feasibility of the concepts introduced above, the author simulated the shop-floor behavior by implementing a prototype of the architecture shown in the Figure 7. Existing MES solutions proved to be inflexible to experiment since they are passive in behavior and hence, do not voluntarily react to the conditions of the shop-floor and also percolate the changes in the PO to GeSCo. In general, the MES solutions provide the directions to the events of the shop-floor only when the information is sought.

In order to engineer a seamless change in PO and to have more control over the simulation, a CMES was developed that mocks the real CMES in the context of production planning and execution (see Figure 8). The SAP Plant Connectivity (SAP-PCo) [44], which is a framework of set of services and management tools was chosen as a basis for GeSCo.

SOAP, REST and an ODATA based web servers, and OPC-UA servers were implemented inside the PCo. During the

research, the PCo was architecturally enhanced to accommodate all the modules of GeSCo (refer Figure 7). These modules were developed inside a Dynamically Linked Library (DLL) along with a set of wrapper operations that were exposed as both web service operations and OPC-UA application specific methods (see Figure 9) that contain the production execution logic. This concept is also called the Enhanced Method Processing (EMP) [45]. The EMP concept enables to cache the production control and routing data, and also embed the orchestration plan algorithms. Furthermore, the EMP implementation assists in behavior specification of the edge component by allowing flexible definition of the actions that need to be executed when invoked by web/OPC-UA client. The EMP DLL is implemented independently by inheriting the API class of the PCo and freely configure the actions that need to be executed during the production. This DLL is imported into the PCo agent instance at design time and the resulting loaded operations/methods are hooked onto the PCo SOAP/REST/ODATA Webserver(s) and/or OPC-UA server(s).

The shop-floor is simulated via a series of Raspberry Pi3 units that act as resources that receive the control instructions from the PCo during production. The experimental simulation setup with reference to the running use case presented in the paper (refer Figure 4) is shown in Figure 10. The simulation

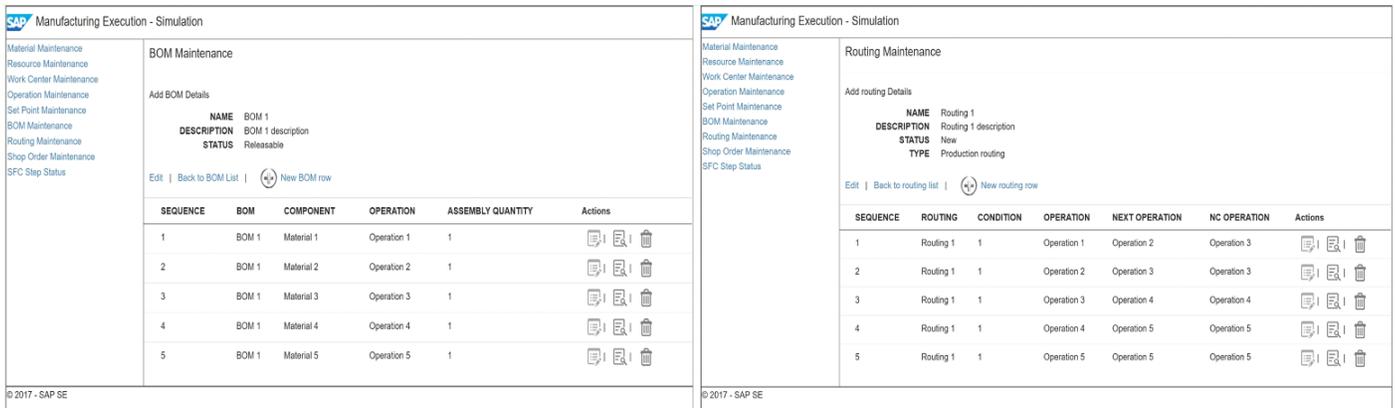


Figure 8: Screen shots of mock CMES user interfaces.

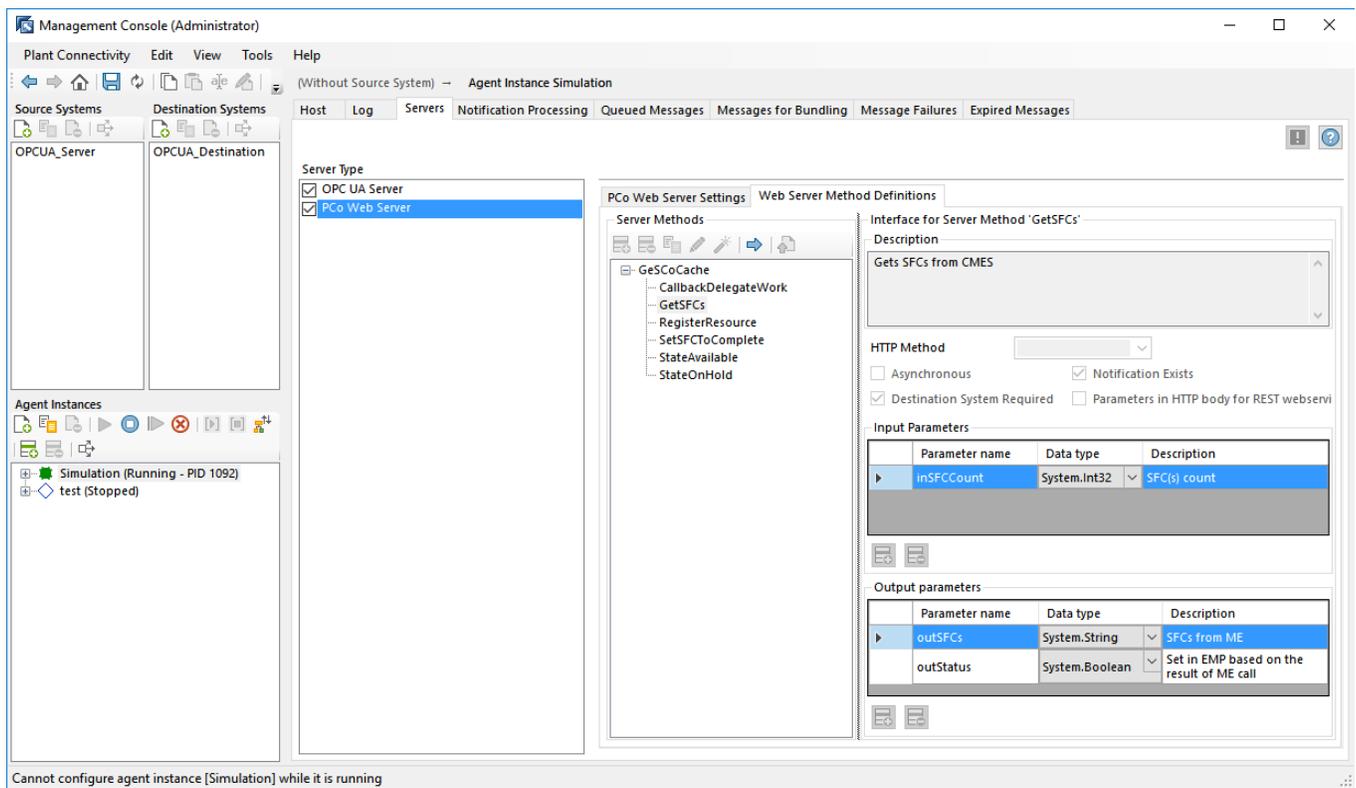


Figure 9: Caching concept implemented as Enhanced Method Processing in PCo.

manufacturing resources supported two important industrial communication protocols: REST based WS-* and OPC-UA. These manufacturing resources acted both as WS-* and OPC-UA client and servers. The capabilities of the manufacturing resources were exposed as a set of web server operations in case of web server and application specific methods in case of OPC-UA server. During the dispatch of the manufacturing operations to the manufacturing resources by the GeSCo, the manufacturing resources act as web or OPC-UA servers. These manufacturing resources operate as WS-* or OPC-UA clients to the GeSCo server during the operation completion acknowledgement step. For the purpose of this simulation, the CMES was geographically separated by approximately

1000km from the GeSCo and mock resource work station deployments to reproduce the typical network latency involved with the cloud solutions, where as the GeSCo and resource work stations were deployed on the same Local Area Network (LAN). A production process without exceptional scenarios that corresponds to the use case illustrated in Figure 4 was simulated to address the challenges 1 and 4 from Section VII-A with different product types of lot size 1, where production routing contained operations that were distributed to resources in a random manner. Two POs with 5 and 3 operations respectively in their routing plan were created in CMES in order to measure the network latency encountered during the production execution. The latency times were measured in the

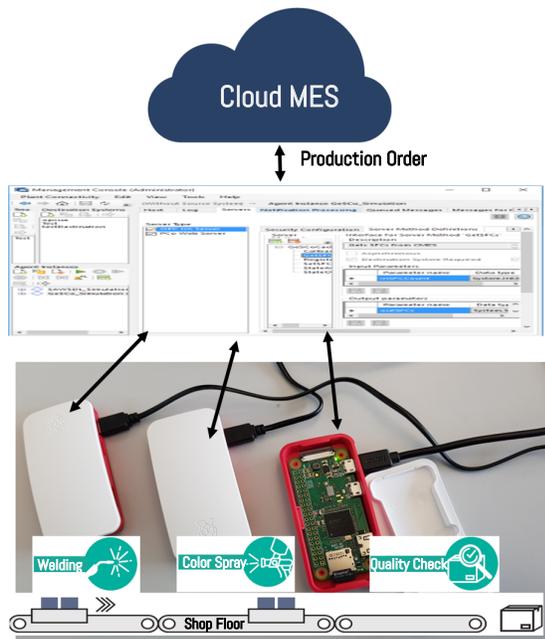


Figure 10: Simulation Setup.



Figure 11: Open Integrated Factory - Generation 2017.

SOAP UI tool [46]. Tables I and II provide the simulation results w.r.t. the network latency encountered without and with GeSCo, respectively. The total latency showed a marked decrease in simulation with the edge layer.

TABLE I: SIMULATION RESULTS WITHOUT GeSCo

Number of Operations in PO	5	3
Client - Server Entities	Resource - CMES	Resource - CMES
Network Latency Per Call	~400 ms	~400 ms
Client - Server calls	10	6
Total Network Latency suffered by PO	~4000 ms	~2400 ms

The research concept was also implemented in the *Open Integrated Factory - Generation 2017* (see Figure 11) that SAP along with other technology partners showcased in *Hannover Industrial Fair - 2017*, which verifies the assumption that the result of simulations is valid under real manufacturing conditions.

TABLE II: SIMULATION RESULTS WITH GeSCo

Number of Operations in PO	5		3	
	GeSCo - CMES	GeSCo - Resource	GeSCo - CMES	GeSCo - Resource
Client - Server Entities	GeSCo - CMES	GeSCo - Resource	GeSCo - CMES	GeSCo - Resource
Network Latency Per Call	~400 ms	~30 ms	~400 ms	~30 ms
Client - Server calls	2	10	2	6
Total Network Latency	~800 ms	~300 ms	~800 ms	~180 ms
Total Network Latency suffered by PO	~1100 ms		~980 ms	

X. CONCLUSION AND FUTURE WORK

This paper describes the various facets of MES solutions and compares the benefits and drawbacks. Based on the arguments, the paper contends that the CMES is better suited in changing production environments than traditional on-premise MES solutions. The most important challenges of cloud solutions are twofold: connectivity and network latency, and security issues. The latter is not the research focus of this work. The former challenge is elucidated and an empirical study is carried out with the aid of an elementary use case. This challenge requires to be addressed in order to make CMES viable in the context of high speed manufacturing.

To overcome the problem of network latency and connectivity associated with CMES, an edge layer called GeSCo that caches the production control data is introduced and a comprehensive architecture is designed to integrate this edge layer with the CMES. The decentralization of the decision-making process in manufacturing was also taken into consideration during the design of the edge component.

Future work includes further refinement in realization of decentralization, development of a semantic data model for GeSCo, research on the extent of caching under given conditions and handling of exceptional scenarios such as quality non-conformance, machine breakdown, priority orders and changes in PO.

ACKNOWLEDGMENT

This research work is supported by a doctoral grant from SAP SE. A non-provisional patent, with ID 15/926,856, on the research of CMES controlled edge manufacturing has been filed in USA on 20-Mar-2018 with the authors as the inventors.

REFERENCES

- [1] B. Katti, M. Schweitzer, and C. Plociennik, "Gesco: Introducing an edge layer between cloud mes and shop-floor in decentralized manufacturing," in 11th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM-2017), Special Session on Edge Computing in factories of the future (ECFoF), 2017, pp. 153–158.
- [2] A. Lenart, "Erp in the cloud—benefits and challenges," Research in systems analysis and design: Models and methods, 2011, pp. 39–50.

- [3] D. Wu, M. J. Greer, D. W. Rosen, and D. Schaefer, "Cloud manufacturing: Strategic vision and state-of-the-art," *Journal of Manufacturing Systems*, vol. 32, no. 4, 2013, pp. 564–579.
- [4] K. D. Bettenhausen and S. Kowalewski, "Cyber-physical systems: Chancen und nutzen aus sicht der automation," *Thesen und Handlungsfelder*; VDI/VDE-Gesellschaft Mess-und Automatisierungstechnik: Düsseldorf, Germany, 2013.
- [5] G. K. Saharidis, Y. Dallery, and F. Karaesmen, "Centralized versus decentralized production planning," *RAIRO-Operations Research*, vol. 40, no. 2, 2006, pp. 113–128.
- [6] C. Anderson and J. J. Bartholdi, "Centralized versus decentralized control in manufacturing: lessons from social insects," *Complexity and complex systems in industry*, 2000, pp. 92–105.
- [7] "Manufacturing Operations Management from Siemens," URL: <https://www.plm.automation.siemens.com/mom/> [Accessed Date: 2018-02-15].
- [8] "IEC 62264-3:2016 Standard," URL: <https://www.iso.org/standard/67480.html> [Accessed Date: 2018-02-15].
- [9] T. H. Luan, L. Gao, Z. Li, Y. Xiang, G. Wei, and L. Sun, "Fog computing: Focusing on mobile users at the edge," *CoRR arXiv preprint arXiv:1502.01815*, 2015.
- [10] K. Ueda, A. Lengyel, and I. Hatono, "Emergent synthesis approaches to control and planning in make to order manufacturing environments," *CIRP Annals-Manufacturing Technology*, vol. 53, no. 1, 2004, pp. 385–388.
- [11] H. Lödding, "Decentralized wip oriented manufacturing control," in *Handbook of Manufacturing Control*. Springer, 2013, pp. 435–452.
- [12] R. Y. Zhong, Q. Dai, T. Qu, G. Hu, and G. Q. Huang, "Rfid-enabled real-time manufacturing execution system for mass-customization production," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 2, 2013, pp. 283–292.
- [13] P. Leitão, "Agent-based distributed manufacturing control: A state-of-the-art survey," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 7, 2009, pp. 979–991.
- [14] A. W. Colombo, R. Schoop, and R. Neubert, "An agent-based intelligent control platform for industrial holonic manufacturing systems," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 1, 2006, pp. 322–337.
- [15] D. Georgakopoulos, P. P. Jayaraman, M. Fazia, M. Villari, and R. Ranjan, "Internet of things and edge cloud computing roadmap for manufacturing," *IEEE Cloud Computing*, vol. 3, no. 4, 2016, pp. 66–73.
- [16] V. Gezer, J. Um, and M. Ruskowski, "An extensible edge computing architecture: Definition, requirements and enablers," in *11th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM-2017), Special Session on Edge Computing in factories of the future (ECFoF)*, 2017, pp. 148–152.
- [17] C. Pach, T. Berger, T. Bonte, and D. Trentesaux, "Orca-fms: a dynamic architecture for the optimized and reactive control of flexible manufacturing scheduling," *Computers in Industry*, vol. 65, no. 4, 2014, pp. 706–720.
- [18] P. Valckenaers, H. Van Brussel, P. Verstraete, B. Saint Germain et al., "Schedule execution in autonomic manufacturing execution systems," *Journal of manufacturing systems*, vol. 26, no. 2, 2007, pp. 75–84.
- [19] B. Montreuil, J.-M. Frayret, and S. D'Amours, "A strategic framework for networked manufacturing," *Computers in industry*, vol. 42, no. 2, 2000, pp. 299–317.
- [20] M. Marques, C. Agostinho, G. Zacharewicz, and R. Jardim-Gonçalves, "Decentralized decision support for intelligent manufacturing in industry 4.0," *Journal of Ambient Intelligence and Smart Environments*, vol. 9, no. 3, 2017, pp. 299–313.
- [21] B. Hubanks, *Self-organizing military logistics*. Cambridge, Mass., 1998.
- [22] D. Mourtzis and M. Doukas, "Decentralized manufacturing systems review: challenges and outlook," *Logistics Research*, vol. 5, no. 3-4, 2012, pp. 113–121.
- [23] D. Wu, M. J. Greer, D. W. Rosen, and D. Schaefer, "Cloud manufacturing: Strategic vision and state-of-the-art," *Journal of Manufacturing Systems*, vol. 32, no. 4, 2013, pp. 564–579.
- [24] F. Tao, Y. Cheng, L. Da Xu, L. Zhang, and B. H. Li, "Cciot-cmfg: cloud computing and internet of things-based cloud manufacturing service system," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, 2014, pp. 1435–1442.
- [25] X. Xu, "From cloud computing to cloud manufacturing," *Robotics and computer-integrated manufacturing*, vol. 28, no. 1, 2012, pp. 75–86.
- [26] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, "Cloud computing the business perspective," *Decision support systems*, vol. 51, no. 1, 2011, pp. 176–189.
- [27] P. Helo, M. Suorsa, Y. Hao, and P. Anussornnitisarn, "Toward a cloud-based manufacturing execution system for distributed manufacturing," *Computers in Industry*, vol. 65, no. 4, 2014, pp. 646–656.
- [28] L. Zhang, H. Guo, F. Tao, Y. Luo, and N. Si, "Flexible management of resource service composition in cloud manufacturing," in *International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2010 IEEE,. IEEE, 2010, pp. 2278–2282.
- [29] J.-J. Hwang, H.-K. Chuang, Y.-C. Hsu, and C.-H. Wu, "A business model for cloud computing based on a separate encryption and decryption service," in *Information Science and Applications (ICISA)*, 2011 International Conference on. IEEE, 2011, pp. 1–7.
- [30] F. Tao, L. Zhang, Y. Liu, Y. Cheng, L. Wang, and X. Xu, "Manufacturing service management in cloud manufacturing: Overview and future research directions," *Journal of Manufacturing Science and Engineering*, vol. 137, no. 4, 2015, p. 040912.
- [31] W. Voorsluys, J. Broberg, and R. Buyya, "Introduction to cloud computing," *Cloud computing: Principles and paradigms*, 2011, pp. 1–41.
- [32] T. Wood, P. J. Shenoy, A. Gerber, J. E. van der Merwe, and K. K. Ramakrishnan, "The case for enterprise-ready virtual private clouds," in *HotCloud*, 2009.
- [33] D. M. Mangiuc, "Enterprise 2.0-is the market ready?" *Accounting and Management Information Systems*, vol. 10, no. 4, 2011, p. 516.
- [34] "CISCO Wiki," URL: https://docwiki.cisco.com/wiki/Introduction_to_WLAN_Technologies [Accessed Date: 2018-02-15].
- [35] R. Ramaswamy, N. Weng, and T. Wolf, "Characterizing network processing delay," in *Global Telecommunications Conference, 2004. GLOBECOM'04*. IEEE, vol. 3. IEEE, 2004, pp. 1629–1634.
- [36] J. F. Kurose and K. W. Ross, "Computer networking: a top-down approach," *Addison Wesley*, vol. 4, 2007, p. 8.
- [37] M. Gagnaire, F. Diaz, C. Coti, C. Cerin, K. Shiozaki, Y. Xu, P. Delort, J.-P. Smets, J. Le Lous, S. Lubiartz et al., "Downtime statistics of current cloud solutions," *International Working Group on Cloud Computing Resiliency*, Tech. Rep, 2014.
- [38] Y.-S. P. Chiu, K.-K. Chen, F.-T. Cheng, and M.-F. Wu, "Optimization of the finite production rate model with scrap, rework and stochastic machine breakdown," *Computers & mathematics with applications*, vol. 59, no. 2, 2010, pp. 919–932.
- [39] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, "About the importance of autonomy and digital twins for the future of manufacturing," *IFAC-PapersOnLine*, vol. 48, no. 3, 2015, pp. 567–572.
- [40] J. Hauptert, "Domeman: Repräsentation, verwaltung und nutzung von digitalen objektgedächtnissen," *Ph.D. dissertation, Saarbrücken, Universität des Saarlandes, Dissertation.*, 2013, 2013.
- [41] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowledge acquisition*, vol. 5, no. 2, 1993, pp. 199–220.
- [42] B. Katti, C. Plociennik, and M. Schweitzer, "SemOPC-UA: Introducing Semantics to OPC-UA Application Specific Methods," in *16th IFAC INCOM 2018 full proceedings*, 2018, p. To Appear.
- [43] B. Katti, C. Plociennik, M. Schweitzer, and M. Ruskowski, "SA-OPC-UA: Introducing Semantics to OPC-UA Application Methods," in *14th IEEE CASE 2018 full proceedings*, 2018, p. To Appear.
- [44] "SAP Plant Connectivity," 2017, URL: https://help.sap.com/viewer/p/SAP_PLANT_CONNECTIVITY [Accessed Date: 2018-02-15].
- [45] "SAP Plant Connectivity, Enhanced Method Processing," 2017, URL: https://help.sap.com/doc/saphelp_mii151sp03/15.1.3/en-US/75/819a57c209ab6be1000000a4450e5/frameset.htm [Accessed Date: 2018-02-15].
- [46] "SOAP UI," 2017, URL: <https://www.soapui.org/> [Accessed Date: 2018-02-15].

Fake Reviews Detection on Movie Reviews through Sentiment Analysis Using Supervised Learning Techniques

Elshrif Elmurngi, Abdelouahed Gherbi
 Department of Software and IT Engineering
 École de Technologie Supérieure
 Montreal, Canada

Email: elshrif.elmurngi.1@ens.etsmtl.ca, abdelouahed.gherbi@etsmtl.ca

Abstract— In recent years, Sentiment Analysis (SA) has become one of the most interesting topics in text analysis, due to its promising commercial benefits. One of the main issues facing SA is how to extract emotions inside the opinion, and how to detect fake positive reviews and fake negative reviews from opinion reviews. Moreover, the opinion reviews obtained from users can be classified into positive or negative reviews, which can be used by a consumer to select a product. This paper aims to classify movie reviews into groups of positive or negative polarity by using machine learning algorithms. In this study, we analyse online movie reviews using SA methods in order to detect fake reviews. SA and text classification methods are applied to a dataset of movie reviews. More specifically, we compare five supervised machine learning algorithms: Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbors (KNN-IBK), KStar (K*) and Decision Tree (DT-J48) for sentiment classification of reviews using three different datasets, including movie review dataset V1.0 and movie reviews dataset V2.0 and movie reviews dataset V3.0. To evaluate the performance of sentiment classification, this work has implemented accuracy, precision, recall and F-measure as a performance measure. The measured results of our experiments show that the SVM algorithm outperforms other algorithms, and that it reaches the highest accuracy not only in text classification, but also in detecting fake reviews.

Keywords- Sentiment Analysis; Fake Reviews; Naïve Bayes; Support Vector Machine; k-Nearest Neighbor; KStar; Decision Tree -J48.

I. INTRODUCTION

Sentiment analysis (SA) is one of the significant domains of machine learning techniques [1]. Opinion Mining (OM), also known as Sentiment Analysis (SA), is the domain of study that analyzes people's opinions, evaluations, sentiments, attitudes, appraisals, and emotions towards entities such as services, individuals, issues, topics, and their attributes [2]. "The sentiment is usually formulated as a two-class classification problem, positive and negative" [2]. Sometimes, time is more precious than money, therefore, instead of spending time in reading and figuring out the positivity or negativity of a review, we can use automated techniques for Sentiment Analysis.

The basis of SA is determining the polarity of a given text at the document, sentence or aspect level, whether the expressed opinion in a document, a sentence or an entity aspect is positive or negative. More specifically, the goals of

SA are to find opinions from reviews and then classify these opinions based upon polarity. According to [3], there are three major classifications in SA, namely: document level, sentence level, and aspect level. Hence, it is important to distinguish between the document level, sentence level, and the aspect level of an analysis process that will determine the different tasks of SA. The document level considers that a document is an opinion on its aspect, and it aims to classify an opinion document as a negative or positive opinion. The sentence level using SA aims to setup opinion stated in every sentence. The aspect level is based on the idea that an opinion consists of a sentiment (positive or negative), and its SA aims to categorize the sentiment based on specific aspects of entities.

The documents used in this work are obtained from a dataset of movie reviews that have been collected by [4] and [10]. Then, an SA technique is applied to classify the documents as real positive and real negative reviews or fake positive and fake negative reviews. Fake negative and fake positive reviews by fraudsters who try to play their competitors existing systems can lead to financial gains for them. This, unfortunately, gives strong incentives to write fake reviews that attempt to intentionally mislead readers by providing unfair reviews to several products for the purpose of damaging their reputation. Detecting such fake reviews is a significant challenge. For example, fake consumer reviews in an e-commerce sector are not only affecting individual consumers but also corrupt purchaser's confidence in online shopping [5]. Our work is mainly directed to SA at the document level, more specifically, on movie reviews dataset. Machine learning techniques and SA methods are expected to have a major positive effect, especially for the detection processes of fake reviews in movie reviews, e-commerce, social commerce environments, and other domains.

In machine learning-based techniques, algorithms such as SVM, NB, and DT-J48 are applied for the classification purposes [6]. SVM is a type of learning algorithm that represents supervised machine learning approaches [7], and it is an excellent successful prediction approach. The SVM is also a robust classification approach [8]. A recent research presented in [3] introduces a survey on different applications and algorithms for SA, but it is only focused on algorithms used in various languages, and the researchers did not focus on detecting fake reviews [9]-[13]. This paper presents five supervised machine learning approaches to classify the sentiment of our dataset, which is compared with two different datasets. We also detect fake positive reviews and

fake negative reviews by using these methods. The main goal of our study is to classify movie reviews as a real reviews or fake reviews using SA algorithms with supervised learning techniques.

The conducted experiments have shown the accuracy, precision, recall, and f-measure of results through sentiment classification algorithms. In three cases (movie reviews dataset V1.0 and movie reviews dataset V2.0 and movie reviews dataset V3.0), we have found that SVM is more accurate than other methods such as NB, KNN-IBK, KStar, and DT-J48.

The main contributions of this study are summarized as follows:

- Using the Weka tool [30], we compare different sentiment classification algorithms, which are used to classify the movie reviews dataset into fake and real reviews.
- We apply the sentiment classification algorithms using three different datasets with stopwords removal. We realized that using the stopwords removal method is more efficient than without stopwords not only in text categorization, but also to detection of fake reviews.
- We perform several analysis and tests to find the learning algorithm in terms of accuracy, precision, recall and F-Measure.

The rest of this paper is organized as follows. Section II presents the related works. Section III shows the methodology. Section IV explains the experiment results, and finally, Section V presents the conclusion and future works.

II. RELATED WORKS

Our study employs statistical methods to evaluate the performance of detection mechanism for fake reviews and evaluate the accuracy of this detection. Hence, we present our literature review on studies that applied statistical methods.

A. Sentiment analysis issues

There are several issues to consider when conducting SA [14]. In this section, two major issues are addressed. First, the viewpoint (or opinion) observed as negative in a situation might be considered positive in another situation. Second, people do not always express opinions in the same way. Most common text processing techniques employ the fact that minor changes between the two text fragments are unlikely to change the actual meaning [14].

B. Textual reviews

Most of the available reputation models depend on numeric data available in different fields; an example is ratings in e-commerce. Also, most of the reputation models focus only on the overall ratings of products without considering the reviews which are provided by customers [15]. On the other hand, most websites allow consumers to

add textual reviews to provide a detailed opinion about the product [16] [17]. These reviews are available for customers to read. Also, customers are increasingly depending on reviews rather than on ratings. Reputation models can use SA methods to extract users' opinions and use this data in the Reputation system. This information may include consumers' opinions about different features [18] and [19].

C. Detecting Fake Reviews Using Machine Learning

Filter and identification of fake reviews have substantial significance [20]. Moraes et al. [21] proposed a technique for categorizing a single topic textual review. A sentiment classified document level is applied for stating a negative or positive sentiment. Supervised learning methods are composed of two phases, namely selection and extraction of reviews utilizing learning models such as SVM.

Extracting the best and most accurate approach and simultaneously categorizing the customers written reviews text into negative or positive opinions has attracted attention as a major research field. Although it is still in an introductory phase, there has been a lot of work related to several languages [22]-[24]. Our work used several supervised learning algorithms such as SVM, NB, KNN-IBK, K* and DT-J48 for Sentiment Classification of text to detect fake reviews.

D. A Comparative Study of different Classification algorithms

Table I shows comparative studies on classification algorithms to verify the best method for detecting fake reviews using different datasets such as News Group dataset, text documents, and movie reviews dataset. It also proves that NB and distributed keyword vectors (DKV) are accurate without detecting fake reviews [12] and [13]. While [11] finds that NB is accurate and a better choice, but it is not oriented for detecting fake reviews. Using the same datasets, [9] finds that SVM is accurate with stopwords method, but it does not focus on detecting fake reviews, while [10] finds that SVM is only accurate without using stopwords method, and also without detecting fake reviews. Sentiment Analysis is a very significant to detect fake reviews [1]. However, they used only supervisor learning techniques based on accuracy and precision. Fundamentally, classification accuracy and precision only are typically not enough information to obtain a good result. However, in our empirical study, results in three cases with movie reviews dataset V1.0 and movie reviews dataset V2.0 and movie reviews dataset V3.0 prove that SVM is robust and accurate for detecting fake reviews by evaluation of measuring the performance with accuracy, precision, F-measure and recall. However, in our empirical study, results in three cases with movie reviews dataset V1.0 and movie reviews dataset V2.0 and movie reviews dataset V3.0 prove that SVM is robust and accurate for detecting fake reviews.

TABLE I. A COMPARATIVE STUDY OF DIFFERENT CLASSIFICATION ALGORITHMS.

Reference	Year	Data Source	Size of dataset	Using Supervised Learning	Language	Classifiers	Detecting Fake Review	Measures	Using stopwords	The best method
[9]	2013	Movie Reviews dataset	2000 Movie Reviews	Yes	English	NB,SVM, kNN	NO	Accuracy, Precision and recall	NO	SVM
[10]	2004	Movie Reviews dataset	2000 Movie Reviews	Yes	English	NB, SVM	NO	Accuracy ,t-test	NO	SVM
[11]	2011	News Group dataset	20 categories with 1000 documents	Yes	English	NB, SVM	NO	Micro-average and macro-average F measure	Yes	NB
[12]	2016	Movie Reviews dataset	4000 movie reviews	Yes	Chinese	NB, SVM, K-NN LLR, Delta TFIDF, LDA-SVM, TFIDF, DKV	NO	precision, recall, F-score as metric, and Accuracy	NO	DKV
[13]	2013	Movie Reviews dataset	1400, 2000 Movie Reviews	Yes	English	NB, SVM	NO	Accuracy, F-measure and Entropy	NO	NB
[1]	2017	Movie Reviews dataset	1400, 2000 Movie Reviews	Yes	English	NB, SVM, IBK, K*,DT-J48	Yes	Precision, and Accuracy	Yes	SVM
This work	2018	Movie Reviews dataset	1400,2000,10662 Movie Reviews	Yes	English	NB, SVM, IBK, K*,DT-J48	Yes	Precision, Accuracy, Recall, and F-Measure	Yes	SVM

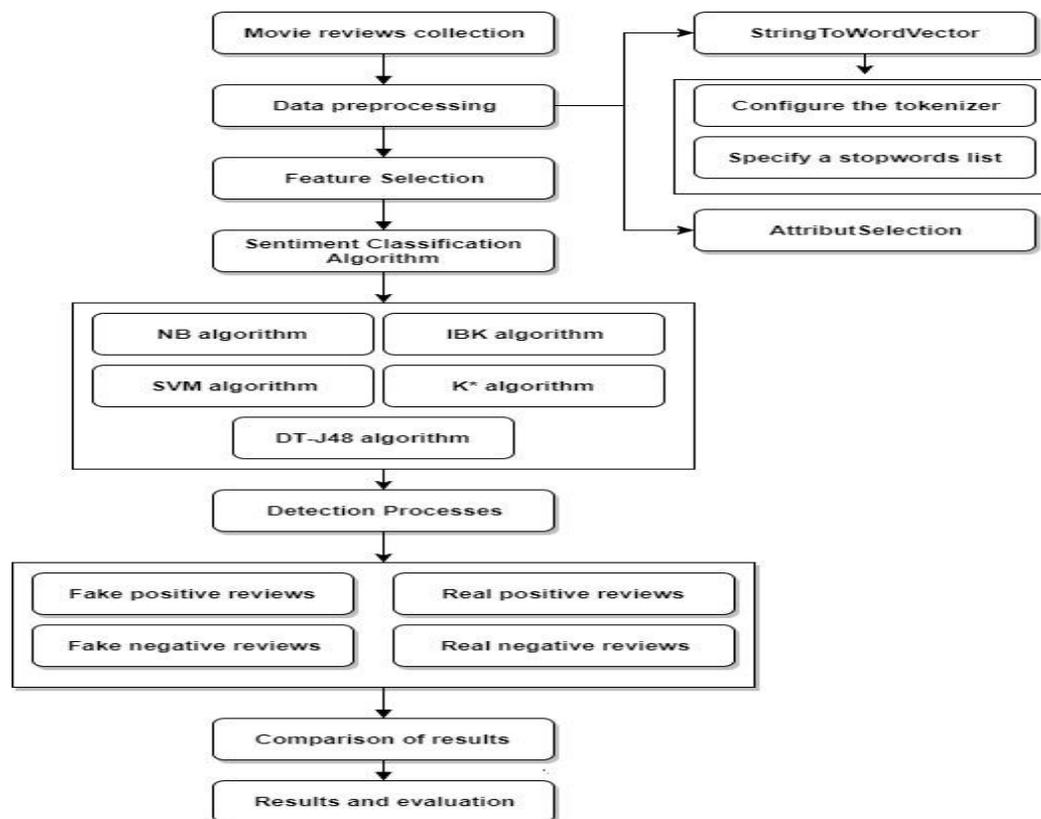


Figure 1. Steps and Techniques used in Sentiment Analysis

III. METHODOLOGY

To accomplish our goal, we analyze a dataset of movie reviews using the Weka tool for text classification. In the proposed methodology, as shown in Figure 1, we follow some steps that are involved in SA using the approaches described below.

Step 1: Movie reviews collection

To provide an exhaustive study of machine learning algorithms, the experiment is based on analyzing the sentiment value of the standard dataset. We have used the original dataset of the movie reviews to test our methods of reviews classification. The dataset is available and has been used in [13], which is frequently conceded as the standard gold dataset for the researchers working in the field of the Sentiment Analysis. The first dataset is known as movie reviews dataset V1.0 which consists of 1400 movie reviews out of which 700 reviews are positive, and 700 reviews are negative. The second dataset is known as movie reviews dataset V2.0, which consists of total 2000 movie reviews, 1000 of which are positive and 1000 of which are negative. The third dataset is known as movie reviews dataset V3.0, which consists of total 10662 movie reviews, 5331 of which are positive and 5331 of which are negative. A summary of the two datasets collected is described in Table II.

TABLE II. DESCRIPTION OF DATASET

Dataset	Content of the Dataset
Movie Reviews Dataset V1.0	1400 Movie Reviews (700+ & 700-)
Movie Reviews Dataset V2.0	2000 Movie Reviews (1000+ & 1000-)
Movie Reviews Dataset V3.0	10662 Movie Reviews (5331+ & 5331-)

Step 2: Data preprocessing

The preprocessing phase includes two preliminary operations, shown in Figure 1, which help in transforming the data before the actual SA task. Data preprocessing plays a significant role in many supervised learning algorithms. We divided data preprocessing as follows:

1) StringToWordVector

To prepare the dataset for learning involves transforming the data by using the StringToWordVector filter, which is the main tool for text analysis in Weka. The StringToWordVector filter makes the attribute value in the transformed datasets Positive or Negative for all single-words, depending on whether the word appears in the document or not. This filtration process is used for configuring the different steps of the term extraction. The filtration process comprises the following two sub-processes:

• Tokenization

This sub-process makes the provided document classifiable by converting the content into a set of features using machine learning.

• Stopwords Removal

The stopwords are the words we want to filter out, eliminate, before training the classifier. Some of those words are commonly used (e.g., "a," "the," "of," "I," "you," "it," "and") but do not give any substantial information to our labeling scheme, but instead they introduce confusion to our classifier. In this study, we used a 630 English stopwords list with movie reviews datasets. Stopwords removal helps to reduce the memory requirements while classifying the reviews.

2) Attribute Selection

Removing the poorly describing attributes can significantly increase the classification accuracy, in order to maintain a better classification accuracy, because not all attributes are relevant to the classification work, and the irrelevant attributes can decrease the performance of the used analysis algorithms, an attribute selection scheme was used for training the classifier.

Step 3: Feature Selection

Feature selection is an approach which is used to identify a subset of features which are mostly related to the target model, and the goal of feature selection is to increase the level of accuracy. In this study, we implemented one feature selection method (BestFirst + CfsSubsetEval, GeneticSearch) widely used for the classification task of SA with Stopwords methods. The results differ from one method to the other. For example, in our analysis of Movie Review datasets, we found that the use of SVM algorithm is proved to be more accurate in the classification task.

Step 4: Sentiment Classification algorithms

In this step, we will use sentiment classification algorithms, and they have been applied in many domains such as commerce, medicine, media, biology, etc. There are many different techniques in classification method like NB, DT-J48, SVM, K-NN, Neural Networks, and Genetic Algorithm. In this study, we will use five popular supervised classifiers: NB, DT-J48, SVM, K-NN, KStar algorithms.

1) Naïve Bayes(NB)

The NB classifier is a basic probabilistic classifier based on applying Bayes' theorem. The NB calculates a set of probabilities by combinations of values in a given dataset. Also, the NB classifier has fast decision-making process.

2) Support Vector Machine (SVM)

SVM in machine learning is a supervised learning model with the related learning algorithm, which examines data and identifies patterns, which is used for regression and classification analysis [25]. Recently, many classification algorithms have been proposed, but SVM is still one of the most widely and most popular used classifiers.

3) *K-Nearest Neighbor (K-NN)*

K-NN is a type of lazy learning algorithm and is a non-parametric approach for categorizing objects based on closest training. The K-NN algorithm is a very simple algorithm for all machine learning. The performance of the K-NN algorithm depends on several different key factors, such as a suitable distance measure, a similarity measure for voting, and, k parameter [26]- [29].

A set of vectors and class labels which are related to each vector constitute each of the training data. In the simplest way; it will be either positive or negative class. In this study, we are using a single number 'k' with values of k=3. This number decides how many neighbors influence the classification.

4) *KStar (K*)*

K-star (K*) is an instance-based classifier. The class of a test instance is established in the class of those training instances similar to it, as decided by some similarity function. K* algorithm is usually slower to evaluate the result.

5) *Decision Tree (DT-J48)*

The DT-J48 approach is useful in the classification problem. In the testing option, we are using percentage split as the preferred method.

Step 5: Detection Processes

After training, the next step is to predict the output of the model on the testing dataset, and then a confusion matrix is generated, which classifies the reviews as positive or negative. The results involve the following attributes:

- True Positive: Real Positive Reviews in the testing data, which are correctly classified by the model as Positive (P).
- False Positive: Fake Positive Reviews in the testing data, which are incorrectly classified by the model as Positive (P).
- True Negative: Real Negative Reviews in the testing data, which are correctly classified by the model as Negative (N).
- False Negative: Fake Negative Reviews in the testing data, which are incorrectly classified by the model as Negative (N).

True negative (TN) are events which are real and are effectively labeled as real, True Positive (TP) are events which are fake and are effectively labeled as fake. Respectively, False Positives (FP) refer to Real events being classified as fakes; False Negatives (FN) are fake events incorrectly classified as Real events. The confusion matrix, (1)-(6) shows numerical parameters that could be applied following measures to evaluate the Detection Process (DP) performance. In Table III, the confusion matrix shows the counts of real and fake predictions obtained with known data, and for each algorithm used in this study there is a different performance evaluation and confusion matrix.

TABLE III. THE CONFUSION MATRIX

	Real	Fake
Real	True Negative Reviews (TN)	False Positive Reviews (FP)
Fake	False Negative Reviews (FN)	True Positive Reviews (TP)

$$\text{Fake Positive Reviews Rate} = \text{FP}/\text{FP}+\text{TN} \quad (1)$$

$$\text{Fake negative Reviews Rate} = \text{FN}/\text{TP}+\text{FN} \quad (2)$$

$$\text{Real Positive Reviews Rate} = \text{TP}/\text{TP}+\text{FN} \quad (3)$$

$$\text{Real negative Reviews Rate} = \text{TN}/\text{TN}+\text{FP} \quad (4)$$

$$\text{Accuracy} = \text{TP}+\text{TN}/\text{TP}+\text{TN}+\text{FN}+\text{FP} \quad (5)$$

$$\text{Precision} = \text{TP}/\text{TP}+\text{FP} \quad (6)$$

$$\text{Recall} = \text{TP}/(\text{TP}+\text{FN}) \quad (7)$$

$$\text{F-measure} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Recall} + \text{Precision}) \quad (8)$$

The confusion matrix is a very important part of our study because we can classify the reviews from datasets whether they are fake or real reviews. The confusion matrix is applied to each of the five algorithms discussed in Step 4.

Step 6: Comparison of results

In this step, we compared the different accuracy provided by the dataset of movie reviews with various classification algorithms and identified the most significant classification algorithm for detecting Fake positive and negative Reviews.

IV. EXPERIMENTS AND RESULT ANALYSIS

In this section, we present experimental results from five different supervised machine learning approaches to classifying sentiment of three datasets which is compared with movie reviews dataset V1.0 and movie reviews dataset V2.0 and movie reviews dataset V3.0. Also, we have used the same methods at the same time to detect fake reviews.

A. *Experimental results on dataset v1.0*

1. Confusion matrix for all methods

The previous section compared different algorithms with different datasets. In this section, the algorithms are applied to perform a sentiment analysis on another dataset. From the results presented in Table IV, the confusion matrix displays results for movie reviews dataset v1.0.

TABLE IV. CONFUSION MATRIX FOR ALL METHODS

Classification algorithms	SA	Real	Fake
NB	Real	455	245
	Fake	162	538
KNN-IBK (K=3)	Real	480	220
	Fake	193	507
K*	Real	491	209
	Fake	219	481
SVM	Real	516	184
	Fake	152	548
DT-J48	Real	498	202
	Fake	219	481

2. Evaluation parameters and accuracy for all methods

Five main performance evaluation measures have been introduced for Classification algorithms. These include Fake Positive Reviews predictive value, Fake Negative Reviews predictive value, Real Positive Reviews predictive value, Real Negative Reviews predictive value, accuracy and Precision. Table V displays the results of evaluation parameters for all methods and provides a summary of recordings obtained from the experiment. As a result, SVM surpasses for best accuracy among the other classification algorithms with 76%.

TABLE V. EVALUATION PARAMETERS AND ACCURACY FOR ALL METHODS

Classification algorithms	Fake Positive Reviews %	Fake Negative Reviews %	Real Positive Reviews %	Real Negative Reviews %	Accuracy %
NB	35	23.1	76.9	65	70.9
K-NN-IBK (K=3)	31.4	27.6	72.4	68.6	70.5
K*	29.9	31.3	68.7	70.1	69.4
SVM	26.3	21.7	78.3	73.7	76
DT-J48	28.9	31.3	68.7	71.1	69.9

The graph in Figure 2 displays a rate of Fake Positive Reviews, Fake Negative Reviews, Real Positive Reviews, Real Negative Reviews, Accuracy for comparative analysis of all different algorithms.

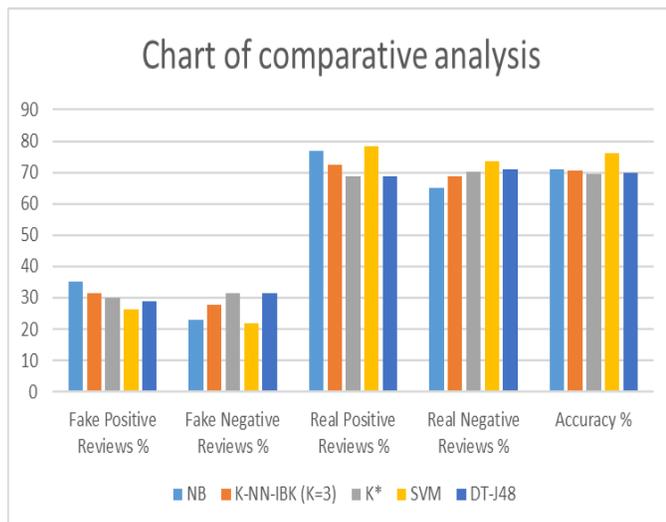


Figure 2. Comparative analysis of all methods

The comparison in Table VI indicates that the classification accuracy of SVM algorithm was better than NB, KNN-IBK, and DT-J48 algorithms.

TABLE VI. COMPARISON OF ACCURACY OF CLASSIFIERS

Classification algorithms	Accuracy %
NB	70.9
KNN-IBK (K=3)	70.5
K*	69.4
SVM	76
DT-J48	69.9

The graph in Figure 3 displays accuracy rate of NB, SVM, (K-NN, k=3), DT-J48 algorithms. We obtained a higher accuracy of SVM algorithm than other algorithms.

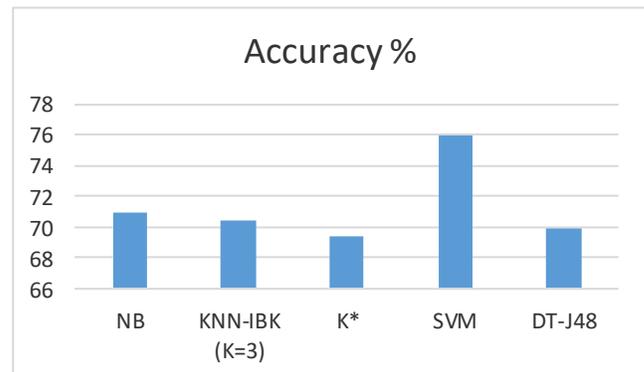


Figure 3. Accuracy of different algorithms

TABLE VII. TIME TAKEN TO BUILD MODEL

Classification algorithms	Time taken to build model (milliseconds)
NB	90
KNN-IBK (K=3)	0
K*	10
SVM	4240
DT-J48	330

Table VII displays the time taken by each algorithm to build prediction model. As it is evident from the table, K-NN takes the shortest amount of time of 0 milliseconds to create a model and SVM takes the longest amount of time of 4240 milliseconds to build a model.

TABLE VIII. COMPARISON RESULTS OF PRECISION, RECALL, AND F-MEASURE

classifier	class	Accuracy metrics %		
		Precision	Recall	F-Measure
NB	pos	68.7	76.9	72.6
	neg	73.7	65.0	69.1
KNN-IBK (K=3)	pos	69.7	72.4	71.1
	neg	71.3	68.6	69.9
K*	pos	69.7	68.7	69.2
	neg	69.2	70.1	69.6
SVM	pos	74.9	78.3	76.5
	neg	77.2	73.7	75.4
DT-J48	pos	70.4	68.7	69.6
	neg	69.5	71.1	70.3

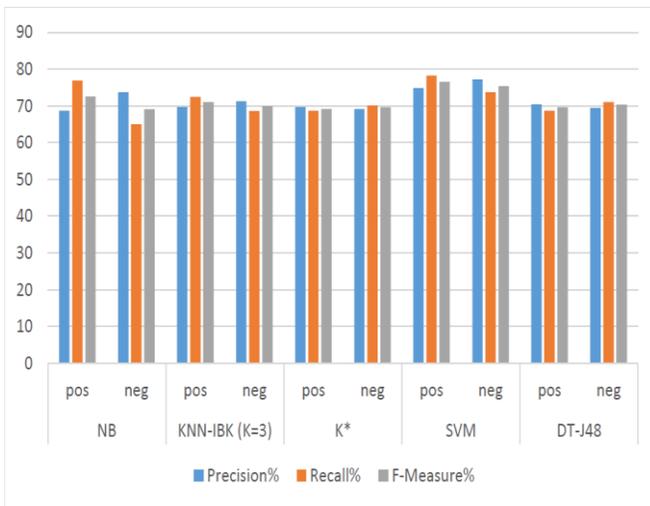


Figure 4. Comparison of metrics obtained from various multi-label classifiers

Table VIII and Figure 4 present the performance evaluation of precision, recall, and f-measure metrics, and all of these metrics are calculated for each class of positive and negative.

B. Experimental result on dataset V2.0

1) Confusion matrix for all methods

The number of real and fake predictions made by the classification model compared with the actual results in the test data is shown in the confusion matrix. The confusion matrix is obtained after implementing NB, SVM, K-NN, K*, DT-J48 algorithms. Table IX displays the results for confusion matrix for V2.0 dataset. The columns represent the number of predicted classifications made by the model. The rows display the number of real classifications in the test data.

TABLE IX. CONFUSION MATRIX FOR ALL METHODS

Classification algorithms	SA	Real	Fake
NB	Real	781	219
	Fake	187	813
KNN-IBK (K=3)	Real	804	196
	Fake	387	613
K*	Real	760	240
	Fake	337	663
SVM	Real	809	191
	Fake	182	818
DT-J48	Real	762	238
	Fake	330	670

2) Evaluation parameters and accuracy for all methods

Five main performance evaluation measures have been introduced for Classification algorithms. These include Fake Positive Reviews predictive value, Fake Negative Reviews predictive value, Real Positive Reviews predictive value, Real Negative Reviews predictive value, accuracy and Precision. Table X shows the results of evaluation parameters for all methods and provides a summary of recordings obtained from the experiment. SVM surpasses as the best accuracy among the other classification algorithms with 81.35%. The tabulated observations list the readings as well as accuracies obtained for a specific supervised learning algorithm on a dataset of a movie review.

TABLE X. EVALUATION PARAMETERS AND ACCURACY FOR ALL METHODS.

Classification algorithms	Fake Positive Reviews %	Fake Negative Reviews %	Real Positive Reviews %	Real Negative Reviews %	Accuracy %
NB	21.9	18.7	81.3	78.1	79.7
K-NN-IBK (K=3)	19.6	38.7	61.3	80.4	70.85
K*	24	33.7	66.3	76	71.15
SVM	19.1	18.2	81.8	80.9	81.35
DT-J48	23.8	33	67	76.2	71.6

The graph in Figure 5 shows a rate of Fake Positive Reviews, Fake Negative Reviews, Real Positive Reviews, Real Negative Reviews, Accuracy for comparative analysis of all different algorithms.

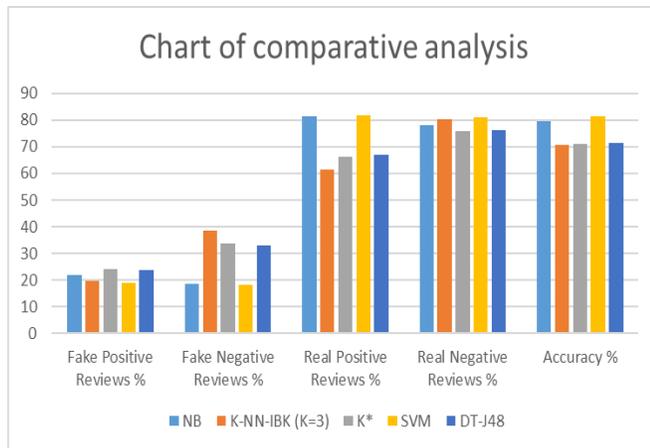


Figure 5. Comparative analysis of all methods

The comparison in Table XI indicates that the classification accuracy of SVM algorithm was better than NB, KNN-IBK, K*, and DT-J48 algorithms.

TABLE XI. COMPARISON OF ACCURACY OF CLASSIFIERS

Classification algorithms	Accuracy %
NB	79.7
KNN-IBK (K=3)	70.85
K*	71.15
SVM	81.35
DT-J48	71.6

The graph in Figure 6 shows accuracy rate of NB, SVM, (K-NN, k=3), and DT-J48 algorithms. We obtained a higher accuracy in SVM algorithm than in the other algorithms.

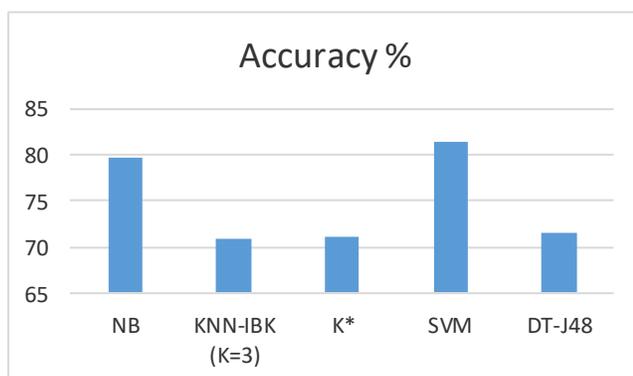


Figure 6. Graph showing the accuracy of different algorithms

Table XII shows the time taken by each algorithm to build prediction model. As it is evident from the table, K-star takes the shortest amount of time of 0 milliseconds to create a model and SVM takes the longest amount of time of **14840** milliseconds to build a model.

TABLE XII. TIME TAKEN TO BUILD MODEL

Classification algorithms	Time taken to build model (milliseconds)
NB	110
KNN-IBK (K=3)	10
K*	0
SVM	14840
DT-J48	340

TABLE XIII. COMPARISON RESULTS OF PRECISION, RECALL, AND F-MEASURE

classifier	class	Accuracy metrics %		
		Precision	Recall	F-Measure
NB	pos	78.8	81.3	80.0
	neg	80.7	78.1	79.4
KNN-IBK (K=3)	pos	75.8	61.3	67.8
	neg	67.5	80.4	73.4
K*	pos	73.4	66.3	69.7
	neg	69.3	76.0	72.5
SVM	pos	81.1	81.8	81.4
	neg	81.6	80.9	81.3
DT-J48	pos	73.8	67.0	70.2
	neg	69.8	76.2	72.8

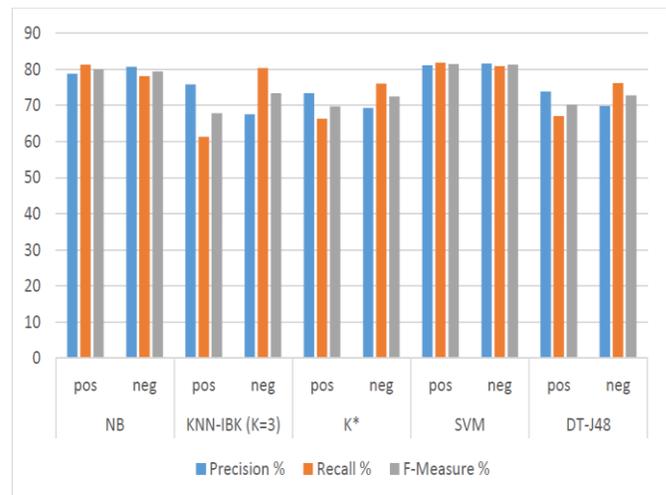


Figure 7. Comparison of metrics obtained from various multi-label classifiers

Table XIII and Figure 7 present the performance evaluation of precision, recall, and f-measure metrics, and all of these metrics are calculated for each class of positive and negative.

C. Experimental results on dataset v3.0

1. Confusion matrix for all methods

The previous section compared different algorithms with different datasets. In this section, the algorithms are applied to perform a sentiment analysis on another dataset. From the results presented in Table XIV, the confusion matrix displays results for movie reviews dataset v3.0.

TABLE XIV. CONFUSION MATRIX FOR ALL METHODS

Classification algorithms	SA		Real	Fake
	Real	Fake		
NB	Real		2303	3028
	Fake		1107	4224
KNN-IBK (K=3)	Real		1813	3518
	Fake		789	4542
K*	Real		2373	2958
	Fake		910	4421
SVM	Real		2758	2573
	Fake		994	4337
DT-J48	Real		2914	2417
	Fake		1571	3760

2. Evaluation parameters and accuracy for all methods

Five main performance evaluation measures have been introduced for Classification algorithms. These include Fake Positive Reviews predictive value, Fake Negative Reviews predictive value, Real Positive Reviews predictive value, Real Negative Reviews predictive value, accuracy and Precision. Table XV displays the results of evaluation parameters for all methods and provides a summary of recordings obtained from the experiment. As a result, SVM surpasses for best accuracy among the other classification algorithms with 66.5%.

TABLE XV. EVALUATION PARAMETERS AND ACCURACY FOR ALL METHODS

Classification algorithms	Fake Positive Reviews %	Fake Negative Reviews %	Real Positive Reviews %	Real Negative Reviews %	Accuracy %
NB	56.8	20.8	79.2	43.2	61.2
K-NN-IBK (K=3)	66	14.8	85.2	34	59.6
K*	55.5	17.1	82.9	44.5	63.7
SVM	48.3	18.6	81.4	51.7	66.5
DT-J48	45.3	29.5	70.5	54.7	62.5

The graph in Figure 8 displays a rate of Fake Positive Reviews, Fake Negative Reviews, Real Positive Reviews, Real Negative Reviews, Accuracy for comparative analysis of all different algorithms.

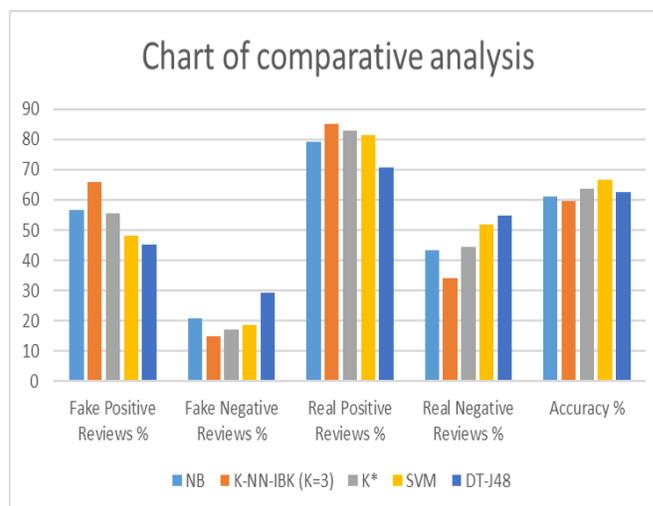


Figure 8. Comparative analysis of all methods

The comparison in Table XVI indicates that the classification accuracy of SVM algorithm was better than NB, KNN-IBK, and DT-J48 algorithms.

TABLE XVI. COMPARISON OF ACCURACY OF CLASSIFIERS

Classification algorithms	Accuracy %
NB	61.2
KNN-IBK (K=3)	59.6
K*	63.7
SVM	66.5
DT-J48	62.5

The graph in Figure 9 displays accuracy rate of NB, SVM, (K-NN, k=3), DT-J48 algorithms. We obtained a higher accuracy of SVM algorithm than other algorithms.

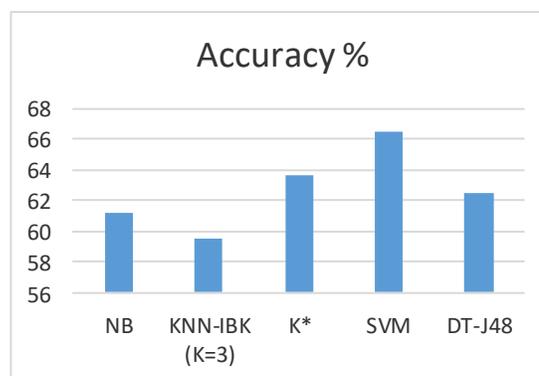


Figure 9. Accuracy of different algorithms

TABLE XVII. TIME TAKEN TO BUILD MODEL

Classification algorithms	Time taken to build model (milliseconds)
NB	680
KNN-IBK (K=3)	20
K*	10
SVM	2,515,260
DT-J48	11,480

Table XVII displays the time taken by each algorithm to build prediction model. As it is evident from the table, K* takes the shortest amount of time of 10 milliseconds to create a model and SVM takes the longest amount of time of 2,515,260 milliseconds to build a model.

TABLE XVIII. COMPARISON RESULTS OF PRECISION, RECALL, AND F-MEASURE

classifier	class	Accuracy metrics %		
		Precision	Recall	F-Measure
NB	pos	58.2	79.2	67.1
	neg	67.5	43.2	52.7
KNN-IBK (K=3)	pos	56.4	85.2	67.8
	neg	69.7	34	45.7
K*	pos	59.9	82.9	69.6
	neg	72.3	44.5	55.1
SVM	pos	62.8	81.4	70.9
	neg	73.5	51.7	60.7
DT-J48	pos	60.9	70.5	65.3
	neg	65	54.7	59.4

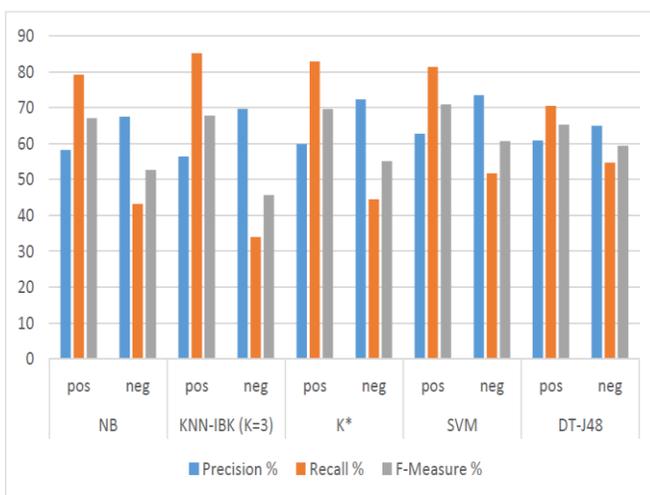


Figure 10. Comparison of metrics obtained from various multi-label classifiers

Table XVIII and Figure 10 present the performance evaluation of precision, recall, and f-measure metrics, and all of these metrics are calculated for each class of positive and negative.

D. Discussion

Table XIX and Figure 11 present the summary of the experiments. Five supervised machine learning algorithms: NB, SVM, K-NN, K*, DT-J48 have been applied to the online movie reviews. We observed that well-trained machine learning algorithms could perform very useful classifications on the sentiment polarities of reviews. In terms of accuracy, SVM is the best algorithm for all tests since it correctly classified 81.35% of the reviews in dataset V1.0 and 76% of the reviews in dataset V2.0 and 66.5% of the reviews in dataset V3.0. SVM tends to be more accurate than other methods.

TABLE XIX. THE BEST RESULT OF OUR EXPERIMENTS

Experiments	Fake Positive Reviews of SVM %	Fake Negative Reviews of SVM %	Accuracy of SVM %
Results on dataset V1.0	19.1	18.2	81.35
Results on dataset V2.0	26.3	21.7	76
Results on dataset V3.0	48.3	18.6	66.5

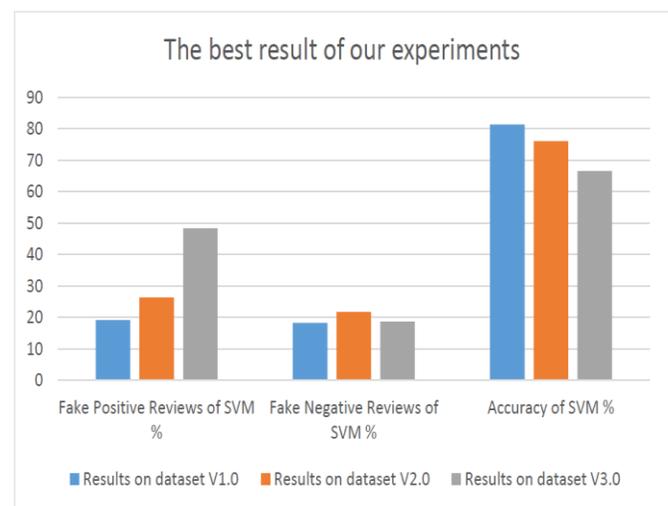


Figure 11. Summary of our experiments

The presented study emphasizes that the accuracy of SVM is higher for Movie Review dataset V2.0. However, the detection process of Fake Positive Reviews and Fake Negative Reviews offers less promising results for Movie Review dataset V2.0 in comparison to Movie Review dataset V1.0 as evident from Table XII.

V. CONCLUSION AND FUTURE WORK

In this research, we proposed several methods to analyze a dataset of movie reviews. We also presented sentiment classification algorithms to apply a supervised learning of the movie reviews located in two different datasets. Our experimental approaches studied the accuracy, precision, recall and F-Measure of all sentiment classification algorithms, and how to determine which algorithm is more accurate. Furthermore, we were able to detect fake positive reviews and fake negative reviews through detection processes.

Five supervised learning algorithms to classifying sentiment of our datasets have been compared in this paper: NB, K-NN, K*, SVM, and DT-J48. Using the accuracy analysis for these five techniques, we found that SVM algorithm is the most accurate for correctly classifying the reviews in movie reviews datasets, i.e., V1.0, V2.0 and V3.0. Also, detection processes for fake positive reviews and fake negative reviews depend on the best method that is used in this study.

For future work, we would like to extend this study to use other datasets such as Amazon dataset or eBay dataset and use different feature selection methods. Furthermore, we may apply sentiment classification algorithms with stopwords removal and stemming methods to detect fake reviews using various tools such as Python or R studio; then we will evaluate the performance of our work with some of these tools.

ACKNOWLEDGMENT

Mr. Elsharif Elmurungi would like to thank the Ministry of Education in Libya and Canadian Bureau for International Education (CBIE) for their support to his Ph.D. research work.

REFERENCES

- [1] E. Elmurungi and A. Gherbi, "Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques," IARIA/ DATA ANALYTICS 2017 – the Sixth International Conference on Data Analytics, ISBN: 978-1-61208-603-3, November, pp. 65–72, Barcelona, Spain, 2017.
- [2] B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, 2012, pp. 1–167.
- [3] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, vol. 5, no. 4, 2014, pp. 1093–1113.
- [4] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," in Proceedings of EMNLP, 2002, pp. 79–86. [Online]. Available: <http://www.cs.cornell.edu/People/pabo/movie%2Dreview%2Ddata/>
- [5] J. Malbon, "Taking fake online consumer reviews seriously," Journal of Consumer Policy, vol. 36, no. 2, 2013, pp. 139–157.
- [6] R. Xia, C. Zong, and S. Li, "Ensemble of feature sets and classification algorithms for sentiment classification," Information Sciences, vol. 181, no. 6, 2011, pp. 1138–1152.
- [7] T. Barbu, "Svm-based human cell detection technique using histograms of oriented gradients," cell, vol. 4, 2012, p. 11.
- [8] G. Esposito, LP-type methods for Optimal Transductive Support Vector Machines. Gennaro Esposito, PhD, 2014, vol. 3.
- [9] P. Kalaivani and K. L. Shunmuganathan, "Sentiment classification of movie reviews by supervised machine learning approaches," Indian Journal of Computer Science and Engineering, vol. 4, no. 4, pp. 285–292, 2013.
- [10] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in Proceedings of the 42nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004, p. 271. [Online]. Available from: <http://www.cs.cornell.edu/People/pabo/movie%2Dreview%2Ddata/>
- [11] S. Hassan, M. Rafi, and M. S. Shaikh, "Comparing svm and naive bayes classifiers for text categorization with wikilogy as knowledge enrichment," in Multitopic Conference (INMIC), 2011 IEEE 14th International. IEEE, 2011, pp. 31–34.
- [12] C.-H. Chu, C.-A. Wang, Y.-C. Chang, Y.-W. Wu, Y.-L. Hsieh, and W.-L. Hsu, "Sentiment analysis on chinese movie review with distributed keyword vector representation," in Technologies and Applications of Artificial Intelligence (TAAI), 2016 Conference on. IEEE, 2016, pp. 84–89.
- [13] V. Singh, R. Piriyani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews and blog posts," in Advance Computing Conference (IACC), 2013 IEEE 3rd International. IEEE, 2013, pp. 893–898.
- [14] G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: a survey," International Journal, vol. 2, no. 6, 2012, pp. 282–292.
- [15] G. Xu, Y. Cao, Y. Zhang, G. Zhang, X. Li, and Z. Feng, "Trm: Computing reputation score by mining reviews," in AAAI Workshop: Incentives and Trust in Electronic Communities, 2016.
- [16] N. Tian, Y. Xu, Y. Li, A. Abdel-Hafez, and A. Josang, "Generating product feature hierarchy from product reviews," in International Conference on Web Information Systems and Technologies. Springer, 2014, pp. 264–278.
- [17] N. Tian, Y. Xu, Y. Li, A. Abdel-Hafez, and A. Josang, "Product feature taxonomy learning based on user reviews," in WEBIST (2), 2014, pp. 184–192.
- [18] A. Abdel-Hafez and Y. Xu, "A survey of user modelling in social media websites," Computer and Information Science, vol. 6, no. 4, 2013, p. 59.
- [19] A. Abdel-Hafez, Y. Xu, and D. Tjondronegoro, "Product reputation model: an opinion mining based approach," in SDAD 2012 The 1st International Workshop on Sentiment Discovery from Affective Data, 2012, p. 16.

- [20] N. Jindal and B. Liu, "Opinion spam and analysis," in Proceedings of the 2008 International Conference on Web Search and Data Mining. ACM, 2008, pp. 219–230.
- [21] R. Moraes, J. F. Valiati, and W. P. G. Neto, "Document-level sentiment classification: An empirical comparison between svm and ann," *Expert Systems with Applications*, vol. 40, no. 2, 2013, pp. 621–633.
- [22] B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web," in Proceedings of the 14th international conference on World Wide Web. ACM, 2005, pp. 342–351.
- [23] A. Fujii and T. Ishikawa, "A system for summarizing and visualizing arguments in subjective documents: Toward supporting decision making," in Proceedings of the Workshop on Sentiment and Subjectivity in Text. Association for Computational Linguistics, 2006, pp. 15–22.
- [24] L.-W. Ku, Y.-T. Liang, and H.-H. Chen, "Opinion extraction, summarization and tracking in news and blog corpora," in Proceedings of AAAI, 2006, pp. 100–107.
- [25] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, 1995.
- [26] Y. Song, J. Huang, D. Zhou, H. Zha, and C. L. Giles, "Iknn: Informativek-nearest neighbor pattern classification," in PKDD. Springer, 2007, pp.248–264.
- [27] G. Bhattacharya, K. Ghosh, and A. S. Chowdhury, "An affinity-based new local distance function and similarity measure for knn algorithm," *Pattern Recognition Letters*, vol. 33, no. 3, 2012, pp. 356–363.
- [28] M. Latourrette, "Toward an explanatory similarity measure for nearest-neighbor classification," *Machine Learning: ECML 2000*, pp. 238–245.
- [29] S. Zhang, "Knn-cf approach: Incorporating certainty factor to knn classification." *IEEE Intelligent Informatics Bulletin*, vol. 11, no. 1, 2010, pp. 24–33.
- [30] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD Explorations newsletter*, vol. 11, no. 1, 2009, pp. 10–18.

Edge Computing and Blockchains for Flexible and Programmable Analytics in Industrial Automation

Mauro Isaja
Research & Development
Engineering Ingegneria Informatica SpA
Rome, Italy
e-mail: mauro.isaja@eng.it

Nikos Kefalakis
IoT Group
Athens Information Technology
Maroussi, Greece
e-mail: nkef@ait.gr

John Soldatos
IoT Group
Athens Information Technology
Maroussi, Greece
e-mail: jsol@ait.gr

Volkan Gezer
Innovative Factory Systems (IFS)
DFKI
Kaiserslautern, Germany
e-mail: Volkan.Gezer@dfki.de

Abstract - The advent of Industry 4.0 has given rise to the introduction of new industrial automation architectures that emphasize the use of digital technologies. In this paper, we present a novel, standards-based Reference Architecture for industrial automation, which combines the benefits of edge computing and blockchain technologies for flexible and reliable orchestration of automation workflows and distributed data analytics. Accordingly, we illustrate a practical implementation of the Reference Architecture for scalable and programmable data analytics, along with its deployment in an Industry 4.0 pilot plant.

Keywords-component; Factory automation; edge computing; blockchain; RAMI4.0; IIRA; Industry4.0; distributed data analytics; programmability; distributed ledger

I. INTRODUCTION

The vision of future manufacturing foresees flexible and hyper-efficient plants that will enable manufacturers to support the transition from conventional “made-to-stock” production models, to the emerging customized ones such as “made-to-order”, “configure-to-order” and “engineering-to-order” [1]. Flexibility in automation is a key prerequisite to supporting the latter production models: It facilitates manufacturers to change automation configurations and rapidly adopt new automation technologies, as a means of supporting variation in production without any essential increase in production costs.

In order to support flexibility in automation, the industrial automation community has been exploring options for the virtualization of the automation pyramid, as part of the transformation of mainstream centralized automation models (like ISA-95) to more distributed ones. Several research and development initiatives have introduced decentralized factory automation solutions based on technologies like intelligent agents [2][3] and Service Oriented Architectures (SOA) [4][5]. These initiatives produced proof-of-concept implementations that highlighted the benefits of

decentralized automation in terms of flexibility. However, they are still not being widely deployed in manufacturing plants, mainly due to the fact that the cost-benefit ratio of such solutions is perceived as unfavourable. Nevertheless, the vision of decentralizing the factory automation pyramid is still alive, as this virtualization can potentially make production systems more agile, increase product quality and reduce cost.

With the advent of the fourth industrial revolution (Industry 4.0) and the Industrial Internet of Things (IIoT), decentralization is being revisited in the light of the integration of Cyber-Physical Systems (CPS) with cloud computing infrastructures. Therefore, several cloud-based applications are deployed and used in factories, which leverage the capacity and scalability of the cloud while fostering supply chain collaboration and virtual manufacturing chains. Early implementations have also revealed the limitations of the cloud in terms of efficient bandwidth usage and its ability to support real-time operations, including operations close to the field.

More recently, the edge computing paradigm has been explored in order to alleviate the limitations of cloud-centric architectures. Edge computing architectures move some part of the system’s overall computing power from the cloud to its edge nodes, i.e., on the field or in close proximity to it – as a means of [6][7]:

- Saving bandwidth and storage, as edge nodes can filter data streams from the field in order to get rid of information without value for industrial automation.
- Enabling low-latency and proximity processing, since information can be processed close to the field.
- Providing enhanced scalability, through supporting decentralized storage and processing that scales better than cloud processing.
- Supporting shopfloor isolation and privacy-friendliness, since edge nodes at the shopfloor are isolated from the rest of the network.

These benefits make edge computing suitable for specific classes of use cases in factories, including:

- Large scale distributed applications, typically applications that involve multiple plants or factories, which process streams from numerous devices at scale.
- Near-real-time applications, which analyse data close to the field or even control CPS systems such as smart machines and industrial robots.

As a result, the application of edge computing to factory automation is extremely promising, since it empowers decentralization in a way that still supports real-time interactions and scalable analytics. Therefore, it is no accident that there are ongoing efforts to provide edge computing implementations for industrial automation in general and factory automation in particular. Furthermore, reference architectures (RAs) for IIoT and industrial automation exist, which highlight the importance of edge computing for compliant implementations. In this article, we present a reference architecture (RA) for factory automation based on edge computing and distributed ledger technology, which has been specified as part of the H2020 FAR-EDGE project [8]. The FAR-EDGE RA specifies some unique features and capabilities, which differentiate them from other on-going implementations of edge computing for factory automation. Most of these unique features concern the exploitation of Distributed Ledger Technology (DLT), today commonly referred to as “blockchain”, as a means of representing and synchronizing automation and data analytics processes based on Smart Contracts. These can be dynamically configured, stored securely and executed in a distributed way, enabling flexibility and scalability in factory automation processes. The implementation of such smart contracts can take advantage of existing distributed ledger platforms. In the scope of FAR-EDGE the popular, open source Hyperledger Fabric provides the foundation for implementing smart contracts and synchronizing distributed processes, as presented in later sections.

Overall, the FAR-EDGE RA combines the power of edge computing for performing operations close to the field, with the reliability and trustworthiness benefits of distributed ledger technologies in terms of the synchronization of distributed processes. This is based on the implementation of a tier of edge nodes where edge functionalities are performed, and its combination with a tier of ledger services that is in charge of plant-wide synchronization of edge nodes. As part of the paper we illustrate the implementation of a Distributed Data Analytics (DDA) platform that adheres to the FAR-EDGE RA, leveraging both edge and ledger tier functionalities. This DDA implementation comes with an additional benefit: it is flexibly extensible and programmable in terms of the definition of edge processing capabilities over field data. Combined with the functionalities of the ledger services, this programmability provides integrators of industrial automation solutions with additional flexibility in implementing distributed analytics systems. The implementations of both the ledger services and the programmable Edge Analytics (EA) services are publicly available as open source software.

Note that the present paper represents a significantly extended version of a conference paper that introduced the FAR-EDGE RA [1]. In particular, this paper includes practical insights on the actual implementation of the edge and ledger tiers of the RA, as part of the DDA platform. It is therefore targeted to researchers and practitioners that might be interested in using the RA and/or its open source implementation in their solutions. It is also destined to the IIoT and Industry4.0 open source community, which is starving for novel, yet practical components for industrial automation and distributed analytics. Note that this paper is a significantly extended version of a conference article of the authors [1].

The paper is structured as follows: Section II, following this introduction, presents state-of-the-art specifications and implementations of the edge computing paradigm for factory automation, as well as the current status in the use of blockchains for industrial applications. It also positions FAR-EDGE against them. Section III introduces the FAR-EDGE RA, from a functional and structural perspective. Section IV presents the prototype implementation of the DDA platform, including its programmability features. Section V is devoted to the presentation of the implementation of ledger services, based on extensions over a permissioned blockchain infrastructure, namely IBM’s Hyperledger Fabric. Section VI illustrates the deployment of the EA systems in an Industry4.0 pilot plant. Finally, Section VII concludes the paper.

II. RELATED WORK

Acknowledging the benefits of edge computing for industrial automation, standards development organizations (SDOs) have specified relevant RAs, while industrial organizations are already working towards providing tangible edge computing implementations.

SDOs such as the OpenFog Consortium and the Industrial Internet Consortium (IIC) have produced RAs for industrial automation applications. In particular, the RA of the OpenFog Consortium prescribes a high-level architecture for Internet-of-Things (IoT) systems, which covers industrial IoT use cases. On the other hand, the RA of the IIC [9] outlines the structuring principles of systems for industrial applications. The IIC RA prescribes the use of edge computing components and principles for compliant implementations. It addresses a wide range of industrial use cases in multiple sectors, including factory automation. These RAs have been recently released and their reference implementations are still in their early stages.

A reference implementation of the IIC RA’s edge computing functionalities for factory automation is provided as part of IIC’s edge intelligence testbed [10]. This testbed provides a proof-of-concept implementation of edge computing functionalities on the shopfloor. The focus of the testbed is on configurable edge computing environments, which enable the development and testing of systems and algorithms for EA. Moreover, Dell-EMC has recently announced the EdgeX Foundry framework [11], which is a vendor-neutral open source project hosted by the Linux Foundation that builds a common open framework for IIoT

edge computing. The framework is influenced by the above-listed RAs and was recently released. Other vendors are also incorporating support for edge devices and Edge Gateways in their cloud platforms.

FAR-EDGE is uniquely positioned in the landscape of edge computing solutions for factory automation. In particular, the FAR-EDGE architecture is aligned to the IIC RA, while exploiting concepts from other RAs and standards such as the OpenFog RA and RAMI 4.0 (Reference Architecture Model Industry 4.0) [12]. However, FAR-EDGE explores pathways and offers functionalities that are not addressed by other specifications and reference implementations. In particular, it researches the applicability of disruptive key enabling technologies like DLT and Smart Contracts in factory automation. DLT, while being well understood and thoroughly tested in mission-critical areas like digital currencies (e.g., Bitcoin), have never been applied before to industrial systems. This is mainly due to performance concerns about their use, despite their trustworthiness and reliability benefits. However, in literature the merit of DLT for synchronizing distributed processes have been recently acknowledged [13].

FAR-EDGE aims at demonstrating how a pool of specific Ledger Services built on a generic DLT platform can enable decentralized factory automation in an effective, reliable, scalable and secure way. In particular, Ledger Services are responsible for sharing process state and enforcing business rules across the computing nodes of a distributed system, thus permitting virtual automation and analytics processes that span multiple nodes – or, from a bottom-up perspective, autonomous nodes that cooperate to a common goal. This is one of project’s unique contributions, which sets it apart from similar edge computing efforts and provides increased flexibility and reliability.

III. FAR-EDGE REFERENCE ARCHITECTURE

The FAR-EDGE RA is a conceptual framework that drives the design and the implementation of automation platforms based on edge computing and DLT technologies. It is aligned to IIC’s RA concepts and described from two architectural viewpoints: the functional viewpoint and the structural viewpoint, as outlined in following paragraphs. An overall architecture representation that includes all elements is provided in Figure 1.

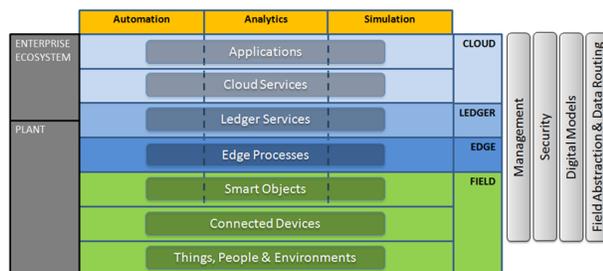


Figure 1. Overview of the FAR-EDGE RA

A. Functional Viewpoint

According to the FAR-EDGE RA, the functionality of a factory automation platform can be decomposed into three high-level Functional Domains - Automation, Analytics and Simulation – and four Crosscutting (XC) Functions – Management, Security, Digital Models and Field Abstraction & Data Routing. To better clarify the scope of such topics, we have tried to map them to similar Industrial Internet RA (IIRA) concepts [9]. Functional Domains and XC Functions are orthogonal to structural Tiers: the implementation of a given functionality may – but is not required to – span multiple Tiers, so that in the overall architecture representation Functional Domains appear as vertical lanes drawn across horizontal layers. Figure 2 highlights the relationship between Functional Domains, their users and the factory environment. It also uses arrows to show the flow of data and of control.

Automation Domain: The FAR-EDGE Automation domain includes functionalities supporting automated control and automated configuration of physical production processes. Automated configuration is the enabler of plug-and-play factory equipment (better known as plug-and-produce), which in turn is a key technology for mass-customization, as it allows a faster and less expensive adjustments of the production process. The Automation domain requires a bidirectional monitoring/control communication channel with the Field, typically with low bandwidth but very strict timing requirements. In some advanced scenarios, Automation is controlled – to some extent – by the results of Analytics and/or Simulation. The Automation domain partially maps to the Control domain of the IIRA.

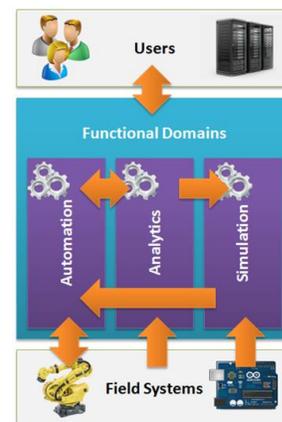


Figure 2. FAR-EDGE RA Functional Domains

Analytics Domain: The FAR-EDGE Analytics domain includes functionalities for gathering and processing Field data for a better understanding of production processes, i.e., a factory-focused business intelligence. This typically requires a high-bandwidth Field communication channel, as the volume of information that needs to be transferred in a given time unit may be substantial. On the other hand, channel latency tends to be less critical than in the Automation scenario. The Analytics domain provides

intelligence to its users, but these are not necessarily limited to humans or vertical applications (e.g., a predictive maintenance solution). In particular, the Automation and Simulation domains, if properly configured, can both make direct use of the outcome of data analysis algorithms. In the case of Automation, the behaviour of a workflow might change in response to changes detected in the controlled process – e.g., a process drift caused by the progressive wear of machinery or by the quality of assembly components being lower than usual. In the case of Simulation, data analysis can be used to update the parameters of a digital model. The Analytics domain matches perfectly the Information domain of the IIRA, except that the latter is receiving data from the Field through the mediation of Control functionalities.

Simulation Domain: The FAR-EDGE Simulation domain includes functionalities for simulating the behaviour of physical production processes for the purpose of optimization or of testing what/if scenarios at minimal cost and risk and without any impact of regular shop activities. Simulation requires digital models of plants and processes to be in-sync with the real-world objects they represent. As the real world is subject to change, models should reflect those changes. For instance, the model of a machine assumes a given value of electric power / energy consumption, but the actual values will diverge as the real machine wears down. To detect this gap and correct the model accordingly, raw data from the Field (direct) or complex analysis algorithms (from Analytics) can be used.

Crosscutting Functions: Crosscutting Functions address common specific concerns. Their implementation affects several Functional Domains and Tiers. They include.

- **Management:** Low-level functions for monitoring and commissioning/decommissioning of individual system modules.
- **Security:** Functions securing the system against the unruly behaviour of its user and of connected systems. These include digital identity management and authentication, access control policy management and enforcement, communication and data encryption.
- **Digital Models:** Functions for the management of digital models and their synchronization with the real-world entities they represent. Digital models are a shared asset, as they may be used as the basis for automated configuration, simulation and field abstraction – e.g., semantic interoperability of heterogeneous field systems.
- **Field Abstraction & Data Routing:** Functions that ensure the connectivity of business logic (FAR-EDGE RA Functional Domains) to the Field, abstracting away the technical details – like device discovery and communication protocols. Data routing refers to the capability of establishing direct producer-consumer channels on demand, optimized for unidirectional massive data streaming – e.g., for feeding Analytics.

B. Structural Viewpoint

The FAR-EDGE RA uses two classes of concepts for describing the structure of a system: Scopes and Tiers.

Scopes are very simple and straightforward: they define a coarse mapping of system elements to either the factory - Plant Scope - or the broader world of corporate IT - Enterprise Ecosystem Scope. Examples of elements in Plant Scope are machinery, Field devices, workstations, SCADA and MES systems, and any software running in the factory data centre. The Enterprise Ecosystem Scope comprises ERP (Enterprise Resource Planning) and PLM (Product Lifecycle Management) systems and any application or service shared across multiple factories or even companies – e.g., supply chain members.

Tiers are a more detailed and technical-oriented classification of deployment concerns. They can be easily mapped to scopes, but they provide more insight into the relationship between system components. This kind of classification is quite similar to OpenFog RA deployment viewpoint, except for the fact that FAR-EDGE Tiers are industry-oriented while OpenFog ones are not. FAR-EDGE Tiers are one of the most innovative traits of the project's RA, and are described in following paragraphs.

The Field Tier is the bottom layer of the FAR-EDGE RA and is populated by Edge Nodes (EN), i.e., any kind of device that is connected to the digital world on one side and to the real world on the other. ENs can have embedded intelligence (e.g., a smart machine) or not (e.g., a sensor or actuator). The FAR-EDGE RA honours this difference: Smart Objects are ENs with on board computing capabilities, Connected Devices are those without. The Smart Object is where local control logic runs: it is a semi-autonomous entity that does not need to interact frequently with the upper layers of the system. As shown in Figure 3. ENs is actually located over field devices.

The Field is also populated by entities of the real world, i.e., those physical elements of production processes that are not directly connected to the network, and as such are not considered as ENs: Things, People and Environments. These are represented in the digital world by some kind of EN wrapper. For instance, room temperature (Environment) is measured by an IoT sensor (Connected Device), the proximity of a worker (People) to a physical checkpoint location is published by an RFID wearable and detected by an RFID Gate (Connected Device), while a conveyor belt (Thing) is operated by a PLC (Smart Object).

The Field Tier is in Plant Scope. Individual ENs are connected to the digital world in the upper Tiers either directly by means of the shopfloor's LAN, or indirectly through some special-purpose local network (e.g., WSN (Wireless Sensor Network)) that is bridged to the former. From the RAMI 4.0 perspective, the FAR-EDGE Field Tier corresponds to the Field Device and Control Device levels on the Hierarchy axis (IEC-62264/IEC-61512), while the entities there contained are positioned across the Asset and Integration Layers.

The Edge Tier is the core of the FAR-EDGE RA. It hosts those parts of Functional Domains and XC Functions

that can leverage the edge computing model, i.e., software designed to run on multiple, distributed computing nodes placed close to the field, which may include resource constrained nodes. The Edge Tier is populated by Edge Gateways (EG): computing devices that act as a digital world gateway to the real world of the Field. These machines are typically more powerful than the average intelligent EN (e.g., blade servers) and are connected to a fast LAN (Local Area Network). Strategically positioned close to physical systems, the EG can execute Edge Processes: time- and bandwidth-critical functionality having local scope. For instance, the orchestration of a complex physical process that is monitored and operated by a number of sensors, actuators (Connected Devices) and embedded controllers (Smart Objects); or the real-time analysis of a huge volume of live data that is streamed from a nearby Field source.

Deploying computing power and data storage in close proximity to where it is actually used is a standard best practice in the industry. However, this technique basically requires that the scope of individual subsystems is narrow (e.g., a single work station). If instead the critical functionality applies to a wider scenario (e.g., an entire plant or enterprise), it must be either deployed at a higher level (e.g., the Cloud) – thus losing all benefits of proximity – or run as multiple parallel instances, each focused on its own narrow scope. In the latter case, new problems may arise: keeping global variables in-sync across all local instances of a given process, reaching a consensus among local instances on a global truth, collecting aggregated results from independent copies of a data analytics algorithm, etc. The need for peer nodes of a distributed system to mutually exchange information is recognized by the OpenFog RA. A key innovation of the FAR-EDGE approach is that it defines a specific system layer – the Ledger Tier – that is responsible for the implementation of such mechanisms and guarantees an appropriate Quality of Service level.

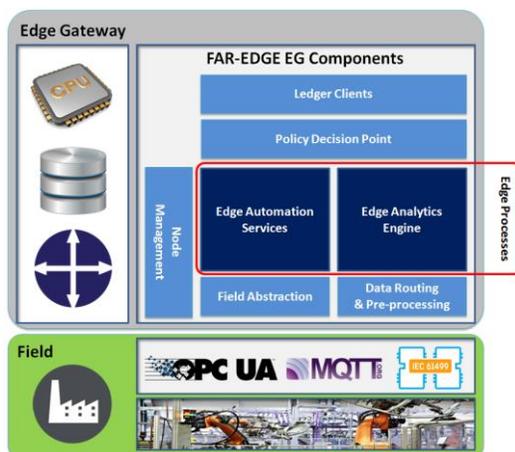


Figure 3. Edge Tier in the FAR-EDGE RA

The Edge Tier is in Plant Scope, located above the Field Tier and below the Cloud Tier. Individual EGs are connected with each other and with the north side of the system, i.e., the globally-scoped digital world in the Cloud Tier – by means

of the factory LAN, and to the south side through the shopfloor LAN. From the RAMI 4.0 perspective, the FAR-EDGE Edge Tier corresponds to the Station and Work Centre levels on the Hierarchy axis (IEC-62264/IEC-61512), while the EGs there contained are positioned across the Asset, Integration and Communication Layers. Edge Processes running on EGs, however, map to the Information and Functional Layers.

The Ledger Tier is a complete abstraction: it does not correspond to any physical deployment environment, and even the entities that it “contains” are abstract. Such entities are Ledger Services, which implement decentralized business logic as smart contracts on top of a distributed ledger. Ledger Services are transaction-oriented: each service call that needs to modify the shared state of a system must be evaluated and approved by Peer Nodes before taking effect. Similarly to “regular” services, Ledger Services are implemented as executable code; however, they are not actually executed on any specific computing node: each service call is executed in parallel by all Peer Nodes that happen to be online at the moment, which then need to reach a consensus on its validity. Most importantly, even the executable code of Ledger Services can be deployed and updated online by means of a distributed ledger transaction.

Ledger Services implement the part of Functional Domains and/or XC Functions that enable the edge computing model, through providing support for their Edge Service counterpart. For example, the Analytics Functional Domain may define a local analytics function (Edge Service) that must be executed in parallel on several EGs, and also a corresponding service call (Ledger Service) that will be invoked from the former each time new or updated local results become available, so that all results can converge into an aggregated data set. In this case, aggregation logic is included in the Ledger Service. Another use case may come from the Automation Functional Domain, demonstrating how the Ledger Tier can also be leveraged from the Field: a smart machine with embedded plug-and-produce functionality can ask permission to join the system by making a service call and then, having received green light, can dynamically deploy its own specific Ledger Service for publishing its state and external high-level commands.

The Ledger Tier lays across the Plant and the Enterprise Ecosystem Scopes, as it can provide support to any Tier. The physical location of Peer Nodes, which implement smart contracts and the distributed ledger, is not defined by the FAR-EDGE RA as it depends on implementation choices.

From the RAMI 4.0 perspective, the FAR-EDGE Ledger Tier corresponds to the Work Centre, Enterprise and Connected World levels on the Hierarchy axis (IEC-62264/IEC-61512), while the Ledger Services are positioned across the Information and Functional Layers.

The Cloud Tier is the top layer of the FAR-EDGE RA, and also the simplest and more “traditional” one. It is populated by Cloud Servers (CS): powerful computing machines, sometimes configured as clusters, which are connected to a fast LAN internally to their hosting data centre, and made accessible from the outside world by means of a corporate LAN or the Internet. On CSs runs that part of

the business logic of Functional Domains and XC Functions that benefits from having the widest of scopes over production processes, and can deal with the downside of being physically deployed far away from them. This includes the planning, monitoring and management of entire factories, enterprises and supply chains (e.g., ERP and SCM (Supply Chain Management) systems). The Cloud Tier is populated by Cloud Services and Applications. Cloud Services implement specialized functions that are provided as individual API calls to Applications, which instead “package” a wider set of related operations that are relevant to some higher-level goal and often expose an interactive human interface.

The Cloud Tier is in Enterprise Ecosystem scope. The “Cloud” term in this context implies that Cloud Services and Applications are visible from all Tiers, wherever located. It does not imply that CSs should be actually hosted on some commercial cloud. In large enterprises, the Cloud Tier corresponds to one or more corporate data centres (private cloud), ensuring that the entire system is fully under the control of its owner.

In terms of RAMI 4.0, the FAR-EDGE Cloud Tier corresponds to the Work Centre, Enterprise and Connected World levels on the Hierarchy axis (IEC-62264/IEC-61512), while the Cloud Services and Applications are positioned across the Information, Functional and Business Layers.

IV. EDGE TIER SERVICES FOR HIGH-PERFORMANCE AND PROGRAMMABLE EDGE ANALYTICS

The FAR-EDGE DDA services span the Edge, Ledger and Cloud Tiers of the FAR-EDGE RA, as illustrated in the following paragraphs.

A. Overview of DDA Tiers

Based on the principles of the FAR-EDGE RA, we have implemented a Distributed Data Analytics (DDA) platform, which enables integrators of factory automation solutions to specify and implement highly distributed data analytics logic, based on data stemming from different parts of a plant. DDA is classified as a reusable, self-sustained component (i.e., “enabler”), which supports the functionalities of the Analytics Domain of the FAR-EDGE RA. The DDA platform implementation spans both the Edge and the Ledger Tiers of the FAR-EDGE RA:

- The Edge Tier that provides the means for accessing and routing field data. Moreover, at the Edge Tier the Edge Analytics Engine (EAE) engine is implemented, which provides the means for executing locally scoped data analytics functionalities, and
- The Ledger Tier leverages “Smart Contracts” that manage analytics configurations. A Smart Contract keeps track and synchronizes information across multiple Edge Gateway nodes. In this way, it provides the means for executing factory-wide data analytics, which span multiple locally scoped analytics functions running in Edge Gateways.

Moreover, the DDA implementation takes advantage of the Cloud Tier as well, where plant-wide data are collected, aggregated and consolidated.

In this section, we present the specification and implementation of the Edge Tier of the DDA platform, which is configurable with almost zero programming. Likewise, the next section illustrates the implementation of the Ledger Services that support the Ledger Tier of the DDA platform.

B. DDA’s Edge Tier: The Edge Analytics Engine (EAE)

The EAE is a runtime environment hosted in an EG, i.e., at the edge of an industrial automation deployment. It is the programmable and configurable environment that executes data analytics logic locally in order to meeting stringent performance requirements, mainly in terms of latency. While the Ledger Services are in charge of managing Smart Contracts and executing distributed analytics across EGs, the EAE is in charge of data analytics within a single EG. The EAE is also configurable, while comprising multiple analytics instances that are driven by multiple smart contracts. It consists of the following main components:

- the EA-Orchestrator;
- the EA-Processor;
- the Local EA-Repository,

which are described in following paragraphs.

The **EA-Orchestrator** provides the run-time environment that controls and executes EA instances, which are specified in a format that is conveniently called Analytics Manifest (AM). In particular, the EA-Orchestrator is able to parse and execute analytics functions and rules specified in an AM. The following statements define the EA-Orchestrator main operation:

- An AM defines a set of EA functionalities, as a graph of processing functions, which can be executed by the EA-Processor.
- The EA-Orchestrator parses an AM and executes the analytics functions that they comprise.
- The EA-Orchestrator is able to execute multiple, concurrent analytics instances. The latter are specified in AMs.

From an implementation perspective, AMs are represented in different forms such as: a configuration file or an entry in a database, or even a part of a smart contract in the blockchain. No matter the implementation technology, the semantics of the AM specify an analytics instance. Hence, the underlying mechanisms that support execution of AMs are independent from specific implementation technologies, as they are based on the implementation agnostic file format that is available as part of the open source implementation of the EAE.

The AM includes the information needed to drive the operation of the EA-Orchestration, including for example the attributes and sequences needed to setup the required jobs on the EA-Processor. As part of its operation the EA-Orchestrator MAY instantiate multiple EA-Processor instances for the purposes of executing an EA instance,

which is described through an AM. Each AM holds the attributes and sequences to set up the required processor jobs in order to serve one EA instance (i.e., one AM).

The **EA-Processor** implements the data processing functionalities that are necessary to implement an EA task. These functionalities are encapsulated in different processor types, including:

- **Pre-processors**, which prepare data streams for analysis, based on the specifications of the target analytics tasks. A pre-processor interacts with a Data Bus in order to acquire streaming data from the field. At the same time, it also produces and registers new streams in the same Data Bus.
- **Analytics Processors**, which apply analytics algorithms to one or more data streams. Similar to the pre-processor, the analytics processor consumes and produces data through interaction with the Data Bus.
- **Store Processors**, which are used to store streams to repositories.

Pre-processors, analytics processors and store processors define three different types of functionalities that are supported by the EAE. Given these processor types, a specific instance of EA is implemented by setting up multiple processors, which are connected in a graph-like fashion thus forming a topology. The topology is specified in the AM, which will be represented as a Smart Contract. The topology and the overall process are controlled by the EA orchestrator.

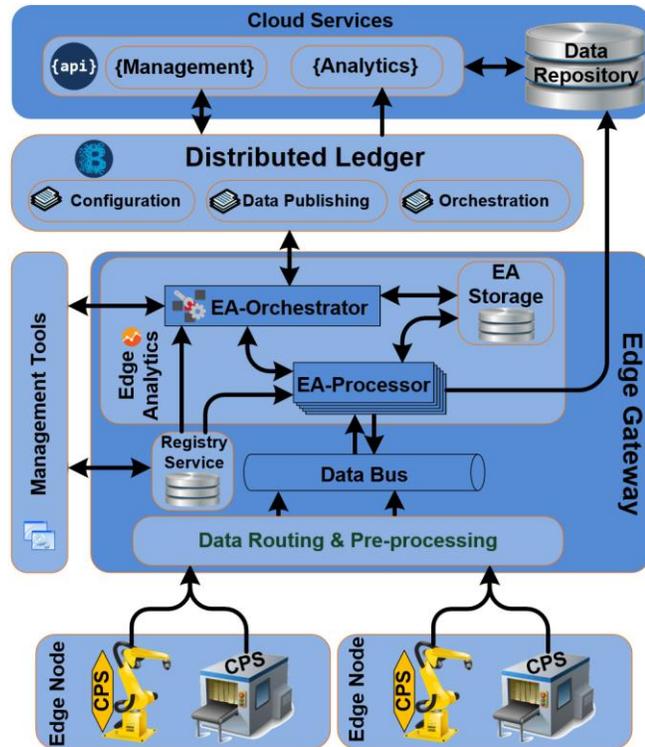


Figure 4. Anatomy of the Edge Analytics Engine

Figure 5 illustrates an example topology and runtime operations for EA Processor. In this example, two streams

(CPS1 and CPS2) are pre-processed from Processor Job 1 (i.e., Pre-Processor) and Processor Job 2 (i.e., Pre-Processor) equivalently in order for an analytics algorithm (i.e., Processor Job 3) to be applied to them. Finally, the result needs to be stored to a Data Storage with the help of Processor Job 4 (i.e., Storage Processor). The setup and runtime operation of the EA-Processor entails the following steps:

- **Step1 (Set-up):** Based on the description of the topology and required processors in the AM, the EA-Orchestrator instantiates and configures the required Processor jobs.
- **Step2 (Runtime):** Processor Job 1 consumes and pre-processes streams coming from CPS1. Likewise, Processor Job 2 consumes and pre-processes streams coming from CPS2.
- **Step3 (Runtime):** Analytics Processor Job 3 consumes the produced streams from Processor Job 1 and 2 for applying the analytics algorithm.
- **Step4 (Runtime):** Store Processor Job 4 consumes the data stream produced from Processor Job 3 and forwards it to the Data Storage.
- **Step5 (Runtime):** Data Storage persists the Data coming from Store Processor Job 4.

Beyond this simple example, much more complex EA workflows can be implemented based on combination of the three different types of processors. The supported scenarios are only limited by the expressiveness of the domain specific language / format that is used to define and represent an AM.

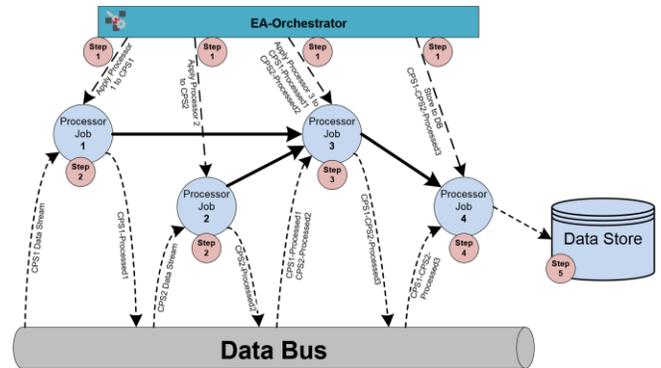


Figure 5. EA Topology Example

C. Using the EAE for Edge Analytics

There are two main ways in which solution developers and integrators can use the EAE:

- **Configuration and execution of analytics queries:** First, they can configure and formulate an analytics query, while they can accordingly execute based on the EAE runtime.
- **Extension of the EAE with analytics capabilities:** Second, they can extend the EAE enabler with additional processing capabilities, which respecting the structure and specification of the engine.

These two ways for taking advantage of the EAE are illustrated in the following paragraphs.

In terms of the configuration and Runtime Execution of Analytics Queries, integrators can take advantage of the EAE API in order to configure and execute analytics queries within an EG. The process includes the following steps:

- **Discovery of Devices:** The first step involves discovery of field devices residing in a devices' registry. Devices define the available data sources to be analyzed by the EAE.
- **Discover available processors:** Following the discovery of devices, available data processors registered in the registry are dynamically identified as well. As already outlined there are three types of processors (i.e., preprocessors, analytics, storage) and multiple instances of each one might be available. Each distinct instance is providing different functionalities based on different implementations.
- **Define and create the Analytics specification:** Based on the available devices and processors, a manufacturer or solution integrator can specify an AM, which defines their desired EA tasks. The definition of the AM comprises a flow of processors, including processor of all three types (i.e., pre-processing, analytics, storage) supported by the EAE engine. It also defines the analytics results to be produced, as well as where they are to be stored / persisted. The specification of the AM can take place based on the use of the EAE's RESTful API. However, in future releases of the EAE we plan to provide a GUI tool in order to facilitate zero-programming specification of the EA tasks.
- **Execute the AM at runtime:** This step involves the runtime execution of the AM through the EA-Orchestrator using its API. With the AM at hand, this step is straightforward and involves the loading an execution of the specification of the manifest. Upon the AM's execution, the analytics results are produced in the forms of name/value pairs, which are stored as specified by the StoreProcessor.

In terms of extending the EAE with Processing & Analytics Capabilities, AMs can be configured and used. AMs provide a convenient mechanism for defining and executing analytics based on a set of available devices and processors. Integrators are able to extend the analytics capabilities of the EAE, based on the specification and deployment of additional processing functions. Additional processing functions have to be of one of the specified types, which will allow their integration and use within AMs.

The process of extended EAE's capabilities involves the following steps:

- **Implementation of a Processor Interface:** In order to extend the EAE with a new processor, an integrator has to provide an implementation of a specific interface, i.e., the interface of the processor. In practice, each of the three processor types comes with its own interface, which specifies its behavior in the scope of the EAE engine.
- **Registration of the Processor to the Registry:** Once a new processor is implemented, it has to become registered to the registry. This will render it discoverable

by solution developers and manufacturers that develop AMs for their needs, based on available devices and processors.

- **Using the processor:** Once a processor becomes available, it can be used for constructing AMs.

D. EAE Open Source Implementation

Apart from a detailed specification of EAE in terms of interfaces, APIs and data schemas for the various processors and the AMs, we have also implemented a prototype of the EAE as open source software [15]. The structure of the implemented system is depicted in Figure 6. As evident in the figure, we take advantage of a Docker container for each distinctive component of our deployment in order to facilitate the distribution, integration and scalability of the system. The Data Bus of the implementation is based on the Apache Kafka platform, which is a distributed system that scales out easily, while offering very high throughput for both publishing and subscribing tasks. Moreover, Kafka supports multi-subscribers and automatically balances the consumers during failures.

The EA-Orchestrator component is also deployed in a Docker container. Hence, the EA-Orchestrator API can be invoked from third party RESTful Client Application (i.e., Postman). To this end, a postman script mapping to the Orchestrator API is offered from GitHub. At the same time, predefined test scripts (i.e., scripts corresponding to AM manifests) have been generated with known actors (CPSs, EA-Processors, configuration attributes etc.).

The EA-Processor component is also deployed in a Docker container. It subscribes to the Data Bus based on the known device IDs. The EA-Processor operates based on a known Number and types of Data Streams. It leverages a static data format.

All available processor types can be used in order to provide a complete test environment including the pre-processing, analytics and analytics storage processors.

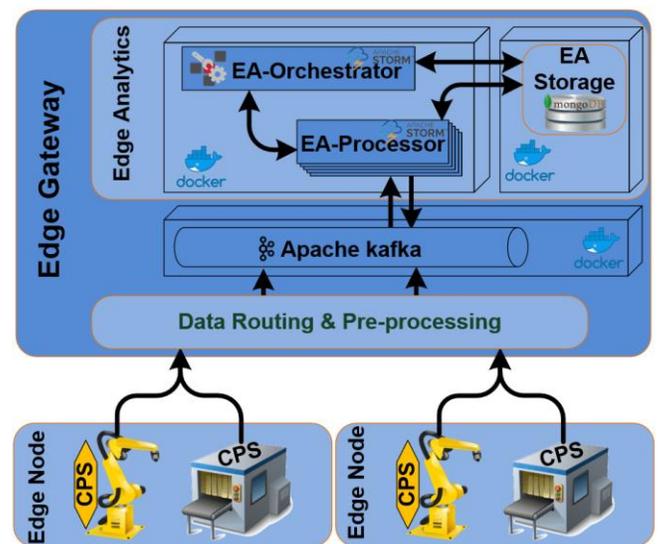


Figure 6. Edge Analytics Engine Implementation

V. LEDGER SERVICES FOR FACTORY WIDE DISTRIBUTED DATA ANALYTICS

The FAR-EDGE Ledger Services enable the most innovative part of the DDA platform, as illustrated in the following paragraphs.

A. Overview

The DDA Platform uses Ledger Services in order to configure plant- and factory-wide analytics processes. Each configuration of analytics algorithms maps to a specific Ledger Service. Every Ledger Service configures one or (usually) more analytics instances. The underlying Distributed Ledger keeps track of multiple analytics configurations. Such configurations are executed by the DDA on production processes that run simultaneously in various locations of the factory.

Moreover, when one analytics task spans multiple EAE instances, a Ledger Service is used to collect local results and implement aggregating logic.

B. Implementation Considerations and Baseline DLT

As explained in the FAR-EDGE RA, the Ledger Tier and Ledger Services are based on DLT – i.e., a Blockchain platform. Concretely, this platform is the Hyperledger Fabric (HLF), which is a commercial-grade Blockchain implementation. HLF has been selected for a number of reasons including its business-friendly open source license, its larger and active community, as well as its support for custom transaction logic (i.e., “smart contracts”) and custom data models. Moreover, HLF is a “permissioned” Blockchain as it supports private networks, which are the primary choice for industrial automation deployments.

The HLF architecture is illustrated in Figure 6. Membership and Orderer are the two elements of the system that are not decentralized, being implemented as central services. Peers, on the contrary, are an arbitrary number of computing nodes that can be deployed anywhere – typically on Edge Gateways – and that run in parallel, providing all the basic services that support the lifecycle of ledger transactions: validation (Endorser), confirmation (Committer), state persistence (Ledger) and listener notification (Events). Last but not least, Peer nodes are where Ledger Services are deployed and run.

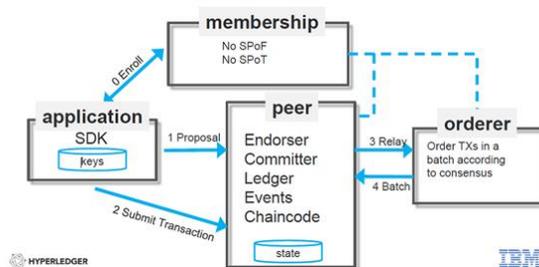


Figure 7. Hyperledger Fabric Logical Architecture [14]

C. Ledger Services

At the platform level, a Ledger Service is a *Chain code* program – i.e., the HLF-specific term for a smart contract. It is designed to support a well-defined, application-specific process. In particular, it is responsible for defining a data model, executing business logic and enforcing access and usage policies. The state of the process is automatically maintained and persisted in the background by the HLF platform, which logs every state change in a distributed ledger that is replicated across all peer nodes.

The data model is shaped by the Chain code itself: a dedicated data store is allocated and initialized by a special code section when the Chain code is first deployed. Once the data store is initialized, no structural changes are expected to happen. It is worth noting that Chaincode instances – and their related data store – are deployed on all peer nodes simultaneously.

Application logic is also coded in the Chain code and is delivered as a number of service endpoints that can be called by clients over the network. These endpoints represent the API of the Ledger Service: only through them callers can query and change its state. The API can be invoked by authorized clients following some well-documented calling conventions of the HLF platform. State-changing calls are managed as a “transaction” by the platform: if the call executes successfully, changes are applied to the persistent storage in all peer nodes; on the other hand, if any error condition is detected (e.g., the Chaincode raising an exception), the platform guarantees that any partial change is reverted.

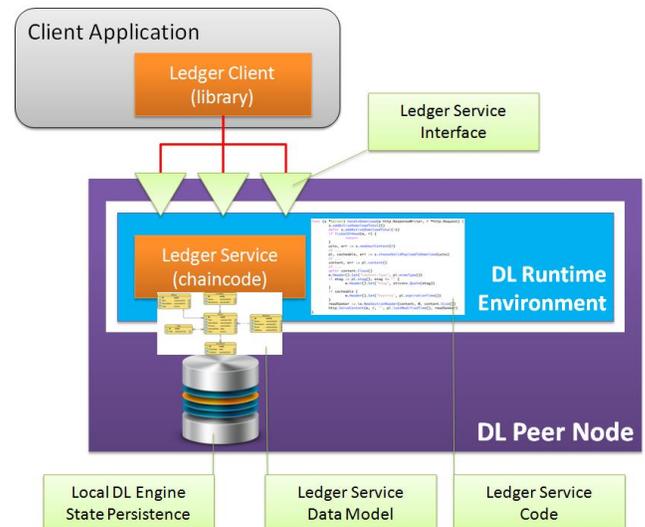


Figure 8. Ledger Services Architecture

In the DDA context, each Ledger Service comes with its own client software library: the Ledger Client. The library provides an in-process API (e.g., Java classes and methods) that matches the network API from a functional point of view but has a much simpler call semantics and hides a lot of HLF-specific technicalities (e.g., user authentication through digital certificates). Ledger Clients can be embedded into

client applications at design time, and used at runtime as a local proxy of the actual Ledger Service API. Figure 8 illustrates the concepts described above from an architectural perspective, focusing on a single peer node.

As already outlined, peer nodes are autonomous sub-systems that run in parallel to provide decentralization and redundancy: each one holds a synchronized copy of the distributed ledger (i.e., the global state of all Ledger Services plus the full history of state-changing transactions) and executes code Ledger Services inside a sandbox environment that isolates each of them from all the others. In order to dynamically adapt the system to the changing needs of the shopfloor, peer nodes can be added to or removed from the running system without any downtime. A number of peer-to-peer protocols are used by peer nodes to collaborate seamlessly with each other, so that the whole system appears to its users as being monolithic. In Figure 9, this relationship is depicted by the “DL Protocols” logical block, that represents the use of common standards for inter-peer communication. Applications can link to the Ledger Service API they are interested in on any peer node of their choice, as all nodes are identical: a service call results in the same code being executed in parallel on each and every node. This redundancy mechanism is what makes the DL a truly decentralized system with exceptional scalability, trustworthiness and reliability properties.

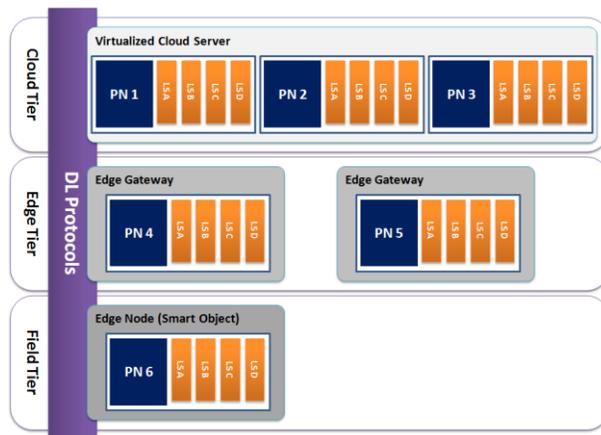


Figure 9. Distributed Ledger Protocol operating across Edge Gateways

In the context of the DDA platform, peer nodes are usually – but not mandatorily – installed on Edge Gateway servers, together with Edge Tier components. This setup allows for clients that run on Edge Gateways, like the EAE, to refer to a local address by default when resolving Ledger Service endpoints. However, peer nodes can be as easily deployed and used on the Cloud Tier, to make them addressable from anywhere; or even embedded into Smart Objects on the Field Tier, to turn the Smart Objects into members of a collaborative P2P (Peer-to-Peer) network.

The definition of access control policies for Ledger Services, and their enforcement at runtime, are built-in features of HLF. There is a fair degree of flexibility in the HLF security subsystem, as individual service endpoints can

be optionally protected by means of attribute-based access control (ABAC). At the most basic level, though, all nodes of the network – including the clients – must have a strong digital identity and be authorized by a central authority in order to join the system. On the other hand, when application-specific control is required, the Ledger Service can manage it as part of the implementation.

D. Self-Adjustment and Reconfiguration (SAR) Service

Self-Adjustment and Reconfiguration (SAR) is an infrastructural feature of the DDA platform. It supports the capability of Smart Objects on the shopfloor to join & leave the system autonomously and to adapt themselves to changing needs and environments in a coordinated way. SAR exploits features of the Ledger Tier, in particular those related to the decentralized coordination of local processes.

The SAR architecture follows the FAR-EDGE RA, spanning three of its layers. The bottom one is the Field Tier, populated by Edge Nodes (EN); right above it, the Edge Tier where a number of Edge Gateways (EG) run some Data Routing components; on top, the Ledger Tier hosting a dedicated Ledger Service: the SAR Service. This design, represented in Figure 10, is driven by a central concept of the FAR-EDGE RA, which breaks down globally-scoped systems into “local clouds”. In the SAR context, Data Routing components on EGs act as “caching proxies” of the SAR Service. More specifically, each EG runs a local device registry that is actually partial view over the master one maintained by the SAR Service. The objective of this design is to allow a local cloud, composed by one EG and a number of EN satellites, to act as a modular unit which can be plugged in and out, and even keep working when temporarily disconnected from the main factory network.

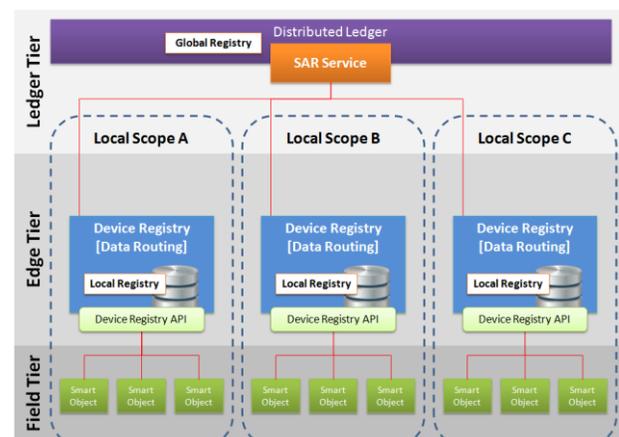


Figure 10. SAR Service Overview

The SAR Service enables the registration, discovery and de-registration of devices that are producers or consumers of *data streams* – i.e., live data flowing from the shopfloor that must be processed in real time. Devices can be either *real* or *virtual*. Real devices can be Smart Objects having the built-in capability of registering and de-registering themselves according to needs (as depicted in Figure 10 above), or

passive IoT sensors that need an administrator to perform these tasks manually. Virtual devices are, instead, computing processes that run on some network node. An example of a virtual data consumer is an analytics program that runs on an EG machine. A virtual data producer may be a program that extracts live data from a legacy database and streams it using some IoT protocol.

The SAR Service also provides endpoints for the creation and decommissioning of communication channels between data producers and consumers. Channels are an abstract notion used to govern how data consumers can connect to data producers, and are the foundation for the enforcement of a device-level access control mechanism. An example use case can help illustrate this point: when a given data consumer wants to establish a new connection to a known data producer (presumably discovered using the registry), it will first need to obtain the authorization to do so from the infrastructure. This is done by means of a SAR Service API call: if successful, this call creates a channel descriptor into the registry. The data producer will then be able to check if incoming connection requests come from “authorized” consumers, and refuse to service them if not.

VI. PILOT PLANT DEPLOYEMENT

We have deployed, tested and demonstrated the DDA in the scope of a pilot plant, which has been built in the scope of the Technology-Initiative Smart Factory-KL. The latter is a testbed for testing and demonstrating the future factory of industrial automation. The plant is arbitrarily modifiable and expandable (flexible), connects arbitrary components of multiple manufacturers (networked), enables its components to perform context-related tasks autonomously (self organizing) and emphasizes user friendliness (user-oriented).

The testbed comprises three Infrastructure Boxes (IB). Each IB comprises energy sensors, which are accessible via an MQTT interface. Energy data are provided every second and comprises information such as the total real power, the total reactive power, the total apparent power, the total real energy, the total reactive energy, the total apparent energy and more. As part of the DDA deployment, we provide the means for computing the hourly daily consumption of the real power and the real energy for each IB and for all three IBs. To this end, on Edge Gateway (comprising an EAE) has been deployed in each one of the IBs.

A data model comprising a Data Interface (DI), a Data Source (DSD) and a Data Kind (DK) has been developed and used to generate a Data Source Manifest (DSM), which is registered in each Edge Gateway. In-line with specification of the EAE, a number of processors have been modelled and developed, including a processor for hourly average calculation from a single data stream, as well as a processor for persisting results in a MongoDB.

The specified at models are used to generate the Analytics Processor Manifest (APM) for each required processor, which is registered to the Edge Gateway. Instances of the above listed processors are created in order

to calculate hourly averages from the total real power and from the total real energy data streams. Moreover, the processor for persisting results is instantiated in order to store results at the edge tier (i.e., in the Edge Gateway’s MongoDB) and at the cloud tier (i.e., a cloud-based MongoDB destined to store global results). The former (edge tier MongoDB) holds the results of EA, while the latter (cloud tier MongoDB) holds the results of factory-wide DDA. Further deployments will be made to get the data from individual Smart Factory-KL modules. These modules can provide additional data such as presence of other nearby modules, current status of the production, state of the module, the order that is being processed along with its priority and other attributes.

Ledger Services are used for orchestrating the instantiated processors. The orchestration is based on an AM, which is registered and controlled through the distributed data Analytics Engine API.

VII. CONCLUSIONS

This paper has introduced a novel RA for decentralized industrial automation, which combines the benefits of edge computing (i.e., near real-time control and data processing) with the capabilities of blockchains in terms of synchronizing distributed processes in scalable way. We have also illustrated a tangible implementation of a DDA platform, which adheres to the main principles of the presented RA. In particular, the DDA platform is empowered by a runtime time environment for programmable, high-performance EA, as well as by a set of distributed ledger services, which enable secure state sharing across multiple analytics processes. The main innovation of the edge tier implementation lies in the fact that it provides the means for specifying, configuring and executing analytics functions with minimal programming. At the same time, the innovation of the ledger services lies in the pioneering use of a permissioned blockchain for synchronizing distributed processes.

The implemented DDA platform and its main enablers are available as open source software [15] [16], which represents one of our tangible contributions to the growing community of Industry4.0 and Industrial IoT researchers and engineers. We have also already deployed a concrete analytics use case in a pilot plant. Our vision and implementation roadmap includes benchmarking the performance of our blockchain system against industry requirements. Based on this benchmarking, we plan to provide to the Industry4.0 community concrete insights on the scope and the limitations of DLT technology for industrial automation and analytics applications.

ACKNOWLEDGMENT

This work has been carried out in the scope of the FAR-EDGE project (H2020-703094). The authors acknowledge help and contributions from all partners of the project.

REFERENCES

- [1] M. Isaja, J. Soldatos, and V. Gezer, "Combining Edge Computing and Blockchains for Flexibility and Performance in Industrial Automation," In the Proc. Of the The Eleventh International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, UBICOMM 2017, November 12 - 16, 2017 - Barcelona, Spain.
- [2] P. Vrba et al., Review of Industrial Applications of Multi-agent Technologies," Service Orientation in Holonic and Multi Agent Manufacturing and Robotics, Studies in Computational Intelligence Vol. 472, Springer, pp 327-338, 2013.
- [3] P. Leitão, "Agent-based distributed manufacturing control: A state-of-the-art survey," Engineering Applications of Artificial Intelligence, vol. 22, no. 7, pp. 979-991, Oct. 2009.
- [4] F. Jammes and H. Smit, "Service-Oriented Paradigms in Industrial Automation Industrial Informatics," IEEE Transactions on, pp. 62 – 70, vol. 1, issue 1, Feb, 2005.
- [5] T. Cucinotta and Coll, "A Real-Time Service-Oriented Architecture for Industrial Automation," Industrial Informatics, IEEE Transactions on, vol. 5, issue 3, pp. 267 – 277, Aug. 2009.
- [6] W. Shi, J. Cao, Q. Zhang, and Y. Li and L. Xu, "Edge Computing: Vision and Challenges," in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, Oct. 2016. doi: 10.1109/JIOT.2016.2579198.
- [7] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," Proceedings of the first edition of the MCC workshop on Mobile cloud computing, MCC '12, pp 13-16.
- [8] H2020-703094 FAR-EDGE Project 2018. [Online], Available from: <http://www.far-edge.eu> 2018.05.30
- [9] Industrial Internet Consortium. 2017. The Industrial Internet of Things Volume G1: Reference Architecture, version 1.8. (2017). [Online], Available from: <http://www.iiconsortium.org/IIRA.htm> 2018.05.30
- [10] Industrial Internet Consortium. IIC Edge Intelligence Testbed. 2017. [Online], Available from: <http://www.iiconsortium.org/edge-intelligence.htm> 2018.05.30
- [11] EdgeX Foundry Framewok 2017. [Online], Available from: <https://www.edgexfoundry.org/> 2018.05.30
- [12] K. Schweichhart. "Reference Architectural Model Industrie 4.0 - An Introduction," April 2016, [Online], Available from: https://ec.europa.eu/futurium/en/system/files/ged/a2-schweichhart-reference_architectural_model_industrie_4.0_rami_4.0.pdf 2018.05.30
- [13] T. Ahram, A. Sargolzaei, S. Sargolzaei, J. Daniels, and B. Amaba, "Blockchain technology innovations," IEEE Technology & Engineering Management Conference (TEMSCON), Santa Clara, CA, 2017, pp. 137-141.
- [14] A. Le Hors and R. Strukhoff, "Hyperledger Fabric Approaches v1.0 with Better Scalability and Security," November 2016. [Online], Available from: <https://www.altoros.com/blog/hyperledger-approaches-version-1-0-with-better-scalability-and-security> 2018.05.30
- [15] FAR-EDGE, EdgeAnalyticsEngine 2018. [Online]. Available from: <https://github.com/far-edge/distributed-data-analytics> 2018.05.30
- [16] FAR-EDGE, DistributedLedger 2018. [Online]. Available from: <https://github.com/far-edge/DistributedLedger> 2018.05.30

Assessing the Application of the In-Memory Technology - A Comprehensive Framework

Stephan Ulbricht*, Marek Opuszko*, Johannes Ruhland*

*Friedrich Schiller University Jena

Department of Business Informations, Jena, Germany

Email: stephan.ulbricht@uni-jena.de, marek.opuszko@uni-jena.de, johannes.ruhland@uni-jena.de

Abstract—After a comeback in recent years, In-memory IT systems are among the most promising solutions to solve present and future IT problems. The use of the In-memory technology promises a massive acceleration of query executions, as demanded by the industry in the face of future challenges such as big data and the internet of things. Despite the increased interest in the technology, however, there is a still hesitant spread. One reason is the lack of practical application scenarios that decision makers can apply to their business context. It has been shown that the sole acceleration of the IT-processing is not sufficient for the dissemination in the business environment. The aim of this work is to introduce a framework to support the evaluation of potential In-memory applications. This design science based framework gives practitioners the opportunity to assess the suitability of In-Memory IT-Systems based on their specific demands. The underlying decision factors have been separated based on their characteristics into value-creation dependent and independent attributes. Appropriate methods were evaluated and selected for the collection of the respective significance. The decision model is implemented using the concept of multi-criteria decision making. The framework is applied using 10 potential real-world In-memory use case scenarios. The results show that the presented approach in this work is suitable to both, assess possible use cases and determine cases with high potential.

Keywords—In-Memory IT-Systems; Business Value; Analytic Hierarchy Process (AHP); In-Memory Technology; In-Memory Computing;

I. INTRODUCTION

In this work, we introduce a design science based framework which reflects both, the industrial as well as the scientific claims to identify and evaluate potential In-memory IT-system (IMIS) scenarios [1]. The decision whether to use an IMIS in a company or not is a complex and multi-criteria decision problem. Because of the fundamentally different approach of IMIS, numerous other aspects beside traditional IT requirements have to be considered. This includes aspects such as, relations to employees, customers and suppliers. Furthermore, possible changes in the company's infrastructure [2] has to be evaluated. The representation of this complexity requires a corresponding model which covers all these different aspects. Due to the versatility of the IMIS technology and its potential use in different use cases, the scenarios may strongly differ. Some aspects may be specific and unique, meaning only relevant for a certain scenario. These aspects are directly linked to the creation of business value and are therefore called value-creation dependent. On the other hand, there will be aspects of a scenario that are not directly linked. These are called value-creation independent. According to their specific characteristics the weightings of the value-creation independent factors are determined by the analytic hierarchy processing and the dependent factors are determined by the direct ranking method. The evaluation and interpretation of the framework is presented based on 10 cross-industry use cases.

Enterprises are faced with the challenge of constantly growing data volumes, increasing competition pressure and the permanent need to instantly react to events. This is one of the main reasons why choosing the "right" IT-systems has become a major strategic decision for companies. The selection of the appropriate system may determine the success of a company or in other words, the selection of the wrong system might lead to serious business disadvantages [3]. The challenges and possibilities associated with the term Big Data characterizes today's IT landscapes. In this context, IMIS represent a key technology [4]. Despite promising expectations, the technology has not yet been significantly established in the industry. Companies mainly criticize the lack of reproducible use cases [5][6]. Since the beginning of the boom of the technology, a whole series of application scenarios have been advertised. Based on these examples, which were often tailored to specific sectors and fields of application, many companies could not derive their own benefits and lead in-memory techniques to fruition. According to a study by the consulting company Pierre Audoin Consultants [7], many companies see great potential in the technology, yet there are only a few cases where the benefits are exploited extensively. This is interesting in contrast to the expectations put on the technology to create business value along all steps of the value chain. This accounts for a vertical as well as a horizontal integration. In addition to these open issues in the corporate sector, there is a clear need for a generalizable reference model to analyze and evaluate in-memory scenarios [8][9] from a scientific perspective. Hence, a universal evaluation tool is needed to determine whether IMIS are beneficial or not suited in a specific scenario and vice versa.

The paper is organized as follows. Section II introduces the research background, the existing literature in the field of IMIS and the overall structure of the framework. Section III presents the research methodology including the analytic hierarchy process (AHP) and the direct ranking method (DRM). In Section IV the application of the framework is shown. The final section summarizes the contributions of this work.

II. RESEARCH BACKGROUND

For a better understanding of the evaluation framework it is necessary to gain a deeper understanding of the technical characteristics of IMIS. The idea of using main memory for the storage of data goes back to the 1980's [10] and 1990's [11]. Caused by the high costs and relatively low storage sizes IMIS were basically a niche technology in the past years. With the introduction of the SAP HANA platform [12], the technology experienced some kind of a comeback. Originally, the SAP HANA platform was developed for accelerated and flexible analysis of large data sets. This new generation of IMIS includes a totally different storage concept in comparison to relational databases. The data in In-Memory systems is

mainly stored in a column-based manner [13]. The advantage is a better data compression [14]. At the same time a column oriented data storage suits better for analytical tasks [15]. The difference between column and row based data storage is shown in Figure 1.

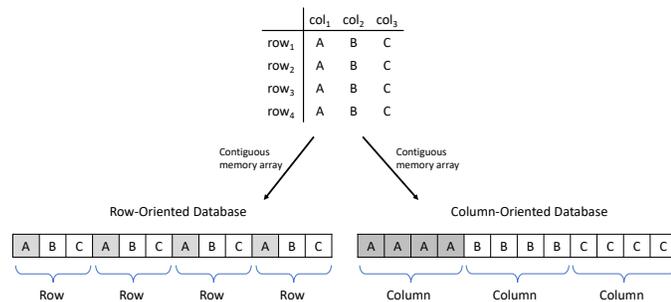


Figure 1. Row- and Column-oriented data layout (adapted according to [15])

In the recent years, the focus on analytical tasks has been extended to hybrid IT-systems. The idea is to store the operational and analytical data entirely in a main memory database [16][17][18]. These hybrid systems are referred to as Online Mixed Workload Processing (OLXP) [19] and Hybrid Transactional/Analytical Processing (HTAP) [20]. Through a common data storage expansive and time consuming extract, transform, load (ETL) processes from the transactional into the analytical system are no longer necessary [15]. As a result, operational data can be used for analysis without major time delays [21].

Due to the different characteristics of analytical and operational tasks, problems and difficulties arise for hybrid systems. The column-based storage of data was originally designed for read-oriented and read-only analysis tasks. A higher proportion of write access typically characterizes operational systems, i.e., enterprise resource planning systems [22], [23]. The merging of these two approaches is often associated with complex join procedures [24][25]. In read-oriented environments, this can reduce the achievable performance improvement promised by IMIS.

A. Problem Context and Related Work

The majority of the early publications in the field of IMIS were characterized by the strong focus on rather technical aspects. To a great proportion, only technical features, such as the column-based storage of data [13], data compression [14] or the persistence of volatile storage media [26] were investigated. The dominance of technical investigations illustrates the strong technologically driven development. Despite its potential, only few studies about the evaluation of IMIS use cases have been published to date. Investigations in the intersection between technology and corporate context have shown that the use of IMIS is not suitable for every application. The speed advantages in comparison to traditional relational databases are related to the number of users and the workload characteristics [27].

The first studies in this field have been carried out by Piller and Hagedorn [8][28][29]. The authors evaluated first case studies in the retail sector. The case studies were evaluated with the aid of various influencing factors. Based on the factors, first application patterns were derived. Another approach to characterize and classify In-memory systems was presented

by Winter et al. [30]. They identified stereotypical patterns based on the data volume and the degree of hybrid workload. An alternative approach for the analysis of In-memory applications addresses the business process characteristics of IMIS use cases. Pioneers in this area were vom Brocke et al. [31][32][33]. They developed a value-creation oriented model, which considers first- as well as second-order effects. They conclude that the value-creation is closely related to process change. The evaluation of several IMIS use cases by Bärenfänger et al. [34] confirmed this results. Cunduis et al. [35][36] considered IMIS from a workflow perspective and developed a framework for the value creation. Another approach focuses on the cost benefit effects of IMIS. In this context, Meier et al. developed a model for the economic evaluation of IMIS. Like vom Brocke et al. they distinguish between direct and indirect benefits. Another work following a cost-benefit perspective can be found in [37].

In their publication, [38] Ulbricht et al. introduce a framework combining the findings of the scientific approaches with practical issues from companies and IMIS system vendors. They presented a structured model for the evaluation and analysis of IMIS use cases, taking various factors into account. Despite the different focuses, one thing all approaches have in common. They all consider the characteristics of IMIS use cases from a quite abstract level. The degree of dissemination in individual sectors, however, indicates the different importance of the particular influencing factors. The question arises, why this technology has already been used quite frequently in some sectors and is hardly ever noticed in other areas.

B. Approach

As mentioned before, the evaluation and analysis of IMIS use cases is a complex, multi-criteria decision problem. In order to represent and solve this problem the concept of the multi-criteria decision making (MCDM) is used. This model allows to consider both, the system requirements as well as the corresponding importance. To determine the total utility U , the weighted sum model (1) of the MCDM [39] is applied. In this model, the system requirements are represented as x_i and the significance (importance) as w_i .

$$U = \sum_{i=1}^n w_i x_i \quad (1)$$

The creation of the framework follows the concept of the design science research [40]. A fundamental requirement of this approach is the proof of the relevance of the developed artifact [41]. The created artifact in this work is represented as a framework. The overall goal of this framework is the decision support regarding the use of the In-memory technology. The relevance of our investigation can be confirmed both from a scientific and a corporate perspective. In order to meet the scientific requirements, the whole design process is based on proven and context-appropriate methods. The applicability of the artifact is demonstrated by real use cases. The overall approach is summarized in Figure 2.

In order to provide a better complexity handling, we characterize the several influence factors and bring them into a hierarchy in the first step. In the second step, we select suitable methods for the determination of the significance depending on the characteristics of the influence factors. The different

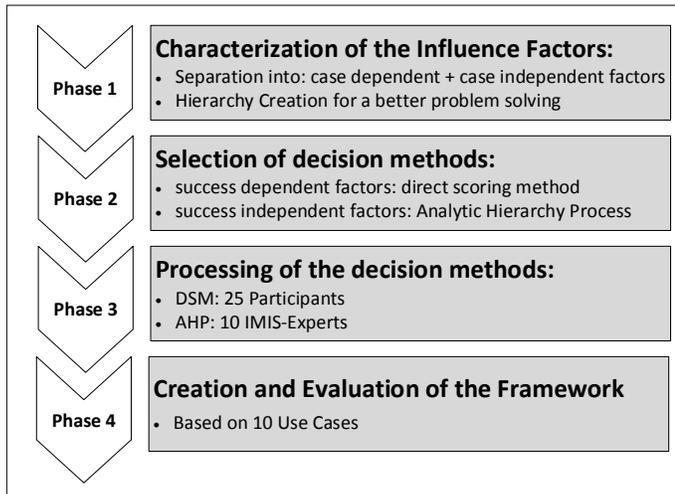


Figure 2. Overview of the Research Methodology

characteristics of the factors lead to a trade off between the operability of the methods and the quality of the results. In the final step, we reveal the results of the utility methods and evaluate the overall framework based on 10 case studies. In this part, we demonstrate the feasibility of our concept by evaluating potential and existing IMIS use cases. Both practical and theoretical aspects are considered in the presented approach. The several steps of the design process are presented in the following sections.

III. RESEARCH METHODOLOGY

After the basic features of the framework have been described in the previous section, the question arises how the respective relevance regarding the evaluation of IMIS is represented. To consider all relevant aspects of an IT investment decision process it is necessary to extend the existing IMIS model by an additional weighting factor. The disjunctive characteristics of the influence factors demand the selection of appropriate weighting methods. Therefore, the decision model was subdivided into a operationalisable goal hierarchy. The categorization as well as the selection process of the corresponding weighting methods is explained in the following part.

A. Characterization and Categorization of the Influence Factors

In [2], DeLone et al. divided the influencing variables of information systems into success dependent and independent. Analogous to this approach we categorized the influence factors in our framework into value-creation dependent and independent. The categorization is presented in the following section. The starting point of the considered influence factors is the IMIS evaluation model from Ulbricht et al. [38]. An overview of the extended framework is given in Figure 3.

Value-creation-dependent influence factors

This category includes the factors, which are crucial for the creation of an economic benefit. Due to the strong impact on the business success, they are particularly important for corporate decisions. These factors comprise the internal as well as the external realization conditions, e.g., the capability to realize the results from the IT-system in an appropriate time. Another influence factor in this category is the potential benefit regarding the use of IMIS. This means value-creation through

faster data processing or more detailed analysis. In most cases, business value is the most important decision criteria for companies. In this consideration, this point also includes non-monetary benefits and second-level effects like an improved customer satisfaction. In addition to these stakeholder-oriented factors, this category also includes technical aspects, which are related to the value-creation. These include, for example, the frequency of change and the range of variation. One of the probably most important advantages of IMIS is the capability of fast data processing. Expert interviews and case studies in this area have shown, however, that the requirements regarding, e.g., the urgency vary significantly between different business areas. In order to achieve independence of the factors, it is important that the potential value generation is considered independent of the other factors. The independence of the decision factors is the prerequisite for the application of methods for the assessment of alternatives[42].

Value-creation-independent influence factors

This category includes factors which are of minor importance from a solely business perspective. This means that these factors have no direct relation to the value-creation. An economically oriented decision maker is in most cases not interested in the underlying data volume or the data structure. On the other hand, these factors play a very important role for the technical evaluation of In-Memory systems. In order to consider all relevant aspects for the evaluation, company representatives, scientists as well as IMIS vendors are involved in the determination of these factors.

B. Method Selection

The determination of weights in the context of information systems have already been the subject of numerous research projects. There exists a plenty of methods to define the relevance of decision alternatives. The challenge for the determination of the weightings is the complexity of the investigated influencing factors in this work. Caused by the distinct characteristics of the factors it is necessary to select appropriate methods for the particular categories. The specification of the value creation dependent weighting factors require the involvement of corporate representatives with a comprehensive understanding of the business processes. As already mentioned in the previous section the significance vary according to the regarded use case. This results in the requirement that the selected method should be easy to use and time-efficient. The value creation independent factors, are determined by technical experts in the field of IMIS. For these factors, however, other properties and corresponding requirements apply. In order to obtain reliable results it is necessary to avoid inconsistencies. At the same time, the method should be proven and well applicable. In the following section the most common procedures are presented and evaluated regarding to their suitability regarding the described requirements.

Direct Rating Method

The direct rating method [43] is a rather straight forward rating method and is applicable to almost every objective target. Given a discrete objective, the direct rating method assigns a value directly to each relevant result x_i of a discrete or nominal target size. In order to keep the evaluation effort manageable, possible alternatives are determined first. In the next step all possible result values are ordered in respect to

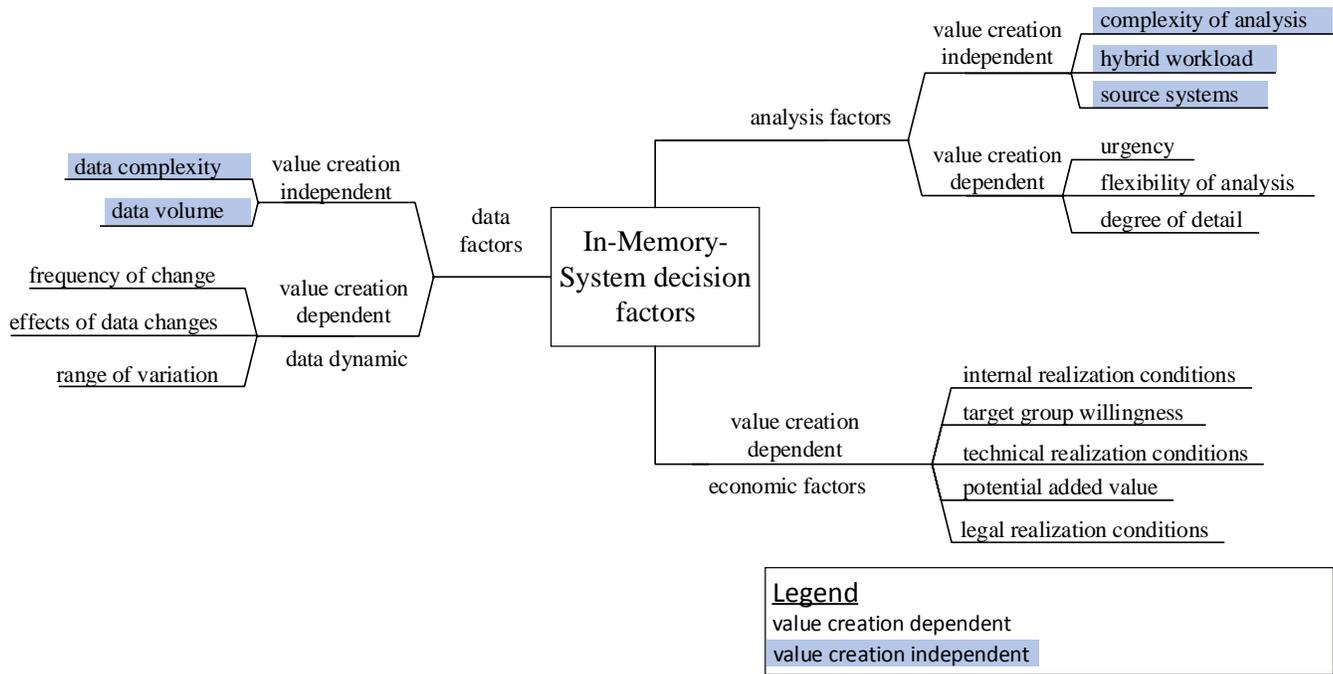


Figure 3. Overview of the analysis and evaluation framework (adapted according to [38])

the given target objective. The lowest result is assigned the lowest value of usually 0 points and the best or highest value is assigned 100 points. For all other result values in between the lowest and highest value, the person evaluating has to assign utility values in order to reflect the pre defined order. There are, however, some issues with this method since there is virtually no support given to the evaluating person on how to assign the actual values. Researchers argue that the direct rating method is more a *backup* method and should only be used in cases of missing alternatives [42]. If a direct assignment is difficult, e.g., for subjective target objectives such as image or taste, proxy attributes should be considered.

Point Allocation Method

In the point allocation method, the decision maker has a given *budget* of points. The points are assigned to the alternatives and reflect the relative importance. For example, the decision maker may be asked to assign 100 points between four alternatives that are relevant to a particular decision. With this method the weights do not have to be normalized since the sum of all assignments has to be 100 points [44]. There is, however, a drawback in the assignment process [45]. In contrast to other methods where the decision maker will assign the value of 100 points to the best alternative and will derive the importance or value of the other alternatives by relating to 100 points, there is no such fixing point in the Point Allocation Method. Especially if there are many options or alternatives, this classification can be difficult. This is because, for example, if the sum of 100 is exceeded, all assessments must be adjusted. In contrast, the assessment of single alternative of the direct rating method has no effect on all other assessments. Jia et al. [46] showed that the point allocation and the direct rating method lead to heterogenous results in terms of the decision weights when applied in practice.

SWING

As part of algebraic methods, the SWING method was

introduced by von Winterfeldt and Edwards [47]. It requires the decision maker to make a cardinal assessments of fully defined (hypothetical) alternatives. In a first step the decision maker defines the null alternative x with all possible attribute values set to the worst possible value. In the next step, possible alternatives $y_i, i = 1..n$ with different attribute values are defined or extracted from the problem context. During a survey the respondent is asked to order all alternatives. The best alternative y_1 is assigned 100 points and the null alternative x is assigned 0 points. Now the respondent is asked to assign points to all other alternatives $y_2..y_n$. In order to determine scale factors, the points of the second best alternative y_2 shall reflect the percentage of the perceived use gain if the respondent *swings* from the null alternative x to the alternative y_i . The best alternative y_1 was already pre-defined with 100 points. Normalized factors are derived through dividing each scale factor by the sum of all scale factors.

To assign attribute weights, the respondent is asked which attribute the respondent would like to change from worst to best, based on the null alternative x . Subsequently, the respondent is asked which attribute should be changed in the next step, until all attributes are assessed. One possible way to determine attribute weights is to set the 100 points to the highest ranked (ordered) attribute and asks the respondent to assign points to all remaining attributes as a scale of 100 points reflecting their importance. Edwards et al. [48] note that it is important to cross check the results with trade off questions, especially when cost is a possible attribute.

Direct Ranking Method

The direct ranking procedure is one of the simplest methods to determine the importance of attributes. At the same time, this method produces the least accurate results of the weight determination methods. In practical environments, the direct ranking is frequently used because of its simple and fast applicability. Compared to other procedures, it is not possible

to check the consistency or plausibility of the answers. The evaluation is carried out by assigning ordinal scaled preference values. In our framework, we use a range from 1 to 10 for the scale. Due to the normalization of the values, the range of the scale is of minor importance. The weighting of the particular factors is obtained by dividing the individual preferences p_i by the total sum of the preferences. The equation for the determination of the weighting is shown in (2).

$$w_i = \frac{p_i}{\sum_{i=1}^n p_i} \quad (2)$$

In spite of the missing methodological variety the direct ranking method suits well for the usage in corporate environments due to its simple applicability. For these reasons, this method was selected for the determination of the value-creation dependent influence factors. To determine the independent parameters more complex methods are necessary.

Analytic Hierarchy Process

The analytic hierarchy process, developed by Saaty [49], is a widely used method for multi-criteria decision problems. This method has been applied in comparable research projects like the selection of enterprise resource planning [50] or the selection of software as a service products [51]. It uses a pairwise comparison of the alternatives to determine ratios and scale priorities. Each factor is compared with every other factor. This kind of comparison improves the decision making within sophisticated problems. On the other hand, with numerous alternatives this leads to an increasing complexity. To reduce this, the alternatives are divided into hierarchies in the AHP. A major advantage with this method is the built-in ability to check the results for inconsistencies. Through the avoidance of inconsistent answers, it is possible to obtain better results in a qualitative manner. However, this requires an increased degree of attention from the participants of a study.

Despite the relatively simple use of pairwise comparisons, the AHP method can produce reliable results. Due to the high complexity and the high demands placed on the participants, this procedure is only suitable to a limited extent for the utilization in companies. This method was selected to determine the weightings of the value-creation independent factors.

IV. APPLICATION OF THE FRAMEWORK

In this section, we present application examples of our IMIS evaluation Framework. In the first part, we determine the weightings of the influence factors, applying the direct ranking method and the AHP. Afterwards, we demonstrate the results of the case studies. For the evaluation of the framework we conducted and analyzed 10 case studies. Thereby, a wide range of companies were involved. This includes, for example, a smaller IT service provider, a medium-sized online travel provider up to a large retailing company. The characteristics of the investigated use cases are shown in Table I. Aimed by the characteristics the evaluation process becomes more comprehensible.

A. Weightings for the Value-Creation Dependent Factors

To determine the business-related significance of the value-creation dependent factors, it was necessary to include only experts with an appropriate extent of knowledge in the field of

data analytics. Therefore, we asked senior corporate representatives in analytic-aware IT positions to rank the importance of each IMIS influence factor. The application of our framework is shown in more detail based on 3 selected use cases. The sample use cases have been chosen considering their business and technical characteristics. In doing so, it is possible to illustrate all aspects of a IMIS use case evaluation. The resulting weightings of the use cases are shown in Table II.

It becomes clear that the significance of the influence factors vary only a bit in the analysis and data categories. Significant differences can be seen within the economic factors. As expected, the potential added value is the most important attribute. Nevertheless, the weightings varies quite strongly. The relatively high influence of the other factors illustrates the need for an overall assessment.

B. Weightings for the Value-Creation Independent Factors

As already mentioned in Section III-B, the mainly technologically driven factors are more complex in their examination. A one-sided investigation from a business perspective does not cover all relevant aspects. It is necessary to involve a broad field of knowledge and experience in this consideration. For this reason, we have included business experts, scientists and experts from system providers to determine these factors. The involved corporate representatives originated from the high-tech industry, the retail sector and online retailing. All these branches are characterized by their high demands regarding the data processing. The representatives have been selected according to their knowledge and experiences in the area of data analytics. The scientific experts were selected from the related work in the context of IMIS. To avoid a technical bias in the investigation the participants from the IMIS system vendors have both, knowledge in the development of such systems but also experiences with the practical application and the customer needs.

A strength of the AHP method is the possibility to detect inconsistent answers. In overall the AHP process includes four phases: the decomposition, the comparative judgements, the determination of priorities and the consistency checking. For a better comprehensibility the proceeding of the AHP as well as an extraction of the results are explained in the following part:

Decomposition

During the decomposition, the underlying decision problem is subdivided into a hierarchical goal system. This step reduces the complexity of the decision-making process. In our example the decomposition is already done through the structuring of the framework into analysis, data and economic factors. Additionally, the elements of the categories were subdivided into value creation dependent and independent.

Comparative Judgements

In the second step the decision criteria are compared pairwise. This comparison is based on a 1 to 9 scale. Where a rating of 1 indicates the equal importance of the considered criteria and a score of 9 indicates the absolute dominance of an attribute over the other. The further meanings of the judgement scale can be found in Table III.

The result of the pairwise comparison is a matrix A with the relative importance of the criterias. In the case of the evaluation

TABLE I. Characteristics of the analyzed Use Cases

		Local Weight Use Case 1	Local Weight Use Case 2	Local Weight Use Case 3	Local Weight Use Case 4	Local Weight Use Case 5	Local Weight Use Case 6	Local Weight Use Case 7	Local Weight Use Case 8	Local Weight Use Case 9	Local Weight Use Case 10
Category	Factor	Analysis of POS-Data	Real-Time Reporting	Finance Reporting	Transaction logging for websites	Analysis of strategic business decisions	Real-time analysis of inventory and sales data	Monitoring and management of IT infrastructure	Project planning	Automated financial accounting	Control of promotions
Analysis	Urgency	Few minutes	Near real-time	Near real-time	few minutes	few hours	near real-time	real time	near real-time	few minutes	real time
	Flexibility of analysis	Ad-hoc	Standard	Standard	Standard	Standard	Standard	-	Standard	Standard	Standard
	Degree of detail	Medium	Very detailed	High	Medium	detailed	detailed	detailed	Very detailed	Very detailed	Medium
	Hybrid workload	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
	Complexity of analysis	High	Very high	Medium	Low	Medium	High	Very low	High	High	Low
	Source systems	2	1	2	2	2	3	2	1	3	2
Data	Data volume	Extremely high	Extremely high	Medium	Medium	Medium	Very high	Very high	Medium	Very high	Very high
	Data complexity	Low	Relatively low	Relatively low	Low	Medium	Relatively low	Relatively low	Medium	Relatively low	Low
	Data dynamic										
	Frequency of change	Rarely	Frequently	Frequently	Frequently	Regularly	Regularly	Very frequently	Frequently	Rarely	Frequently
	Effects of data changes	Low	High	High	Low	Medium	Medium	Medium	Medium	Low	Low
Range of variation	Moderate	Strong changes	Moderate	Little changes	Little changes	Strong changes	Strong changes	Moderate	Moderate	Little changes	
Economic	Internal realization conditions	Months or longer	Hours	Days	Days	Months or longer	Days	Months or longer	Not predictable	Immediate	Immediate
	Potential added value	High	Very high	Medium	Low	High	High	High	Very high	High	Very low
	Target group willingness	Medium	High	Medium	Not predictable	Medium	High	High	Very high	High	Medium
	Technical realization conditions	Low	Low	Medium	Low	High	Low	Not predictable	Low	Low	Very high
	legal realization conditions	Relatively low	Low	High	Not predictable	Relatively low	Relatively low	Relatively low	High	Not predictable	Not predictable

TABLE III. AHP Judgement Scale

Intensity of Importance	Definition
1	Equal Importance
3	Moderate Importance
5	Strong Importance
7	Very strong Importance
9	extreme Importance

TABLE IV. Sample Results of the AHP

	Complexity	Hybrid Workload	Source Systems
Complexity	1	1/5	1/3
Hybrid Workload	5	1	3
Source Systems	3	1/3	1

Determination of Priorities

of IMIS criterias we asked experts in this area to judge the importance of the value creation independent decision factors. A partial result of the assessment can be seen in Table IV. This example shows for instance, that the expert assesses the importance of processing hybrid workload much higher than the complexity of analysis.

For the calculation of the weights the results of the comparison (matrix A) are initially normalized. Thereby, the values are divided by the sum of the respective column values. Formally this relation is shown in (3). In Table V, the results of the calculations are demonstrated based on the continued example. Thereby, it becomes clear that the weightings reflect

TABLE II. WEIGHTINGS OF THE VALUE-CREATION DEPENDENT FACTORS

Category	Factor	Local Weight Use Case 1	Local Weight Use Case 2	Local Weight Use Case 3	Local Weight Use Case 4	Local Weight Use Case 5	Local Weight Use Case 6	Local Weight Use Case 7	Local Weight Use Case 8	Local Weight Use Case 9	Local Weight Use Case 10
Analysis	Urgency	0.306	0.316	0.304	0.417	0.267	0.305	0.414	0.317	0.313	0.407
	Flexibility of analysis	0.421	0.367	0.353	0.322	0.376	0.39	0.401	0.366	0.374	0.389
	Degree of detail	0.272	0.316	0.342	0.261	0.357	0.305	0.184	0.317	0.313	0.204
Data	Data dynamic										
	Frequency of change	0.286	0.333	0.300	0.366	0.331	0.398	0.331	0.426	0.452	0.452
	Effects of data changes	0.286	0.333	0.400	0.513	0.541	0.487	0.541	0.464	0.443	0.443
	Range of variation	0.429	0.333	0.300	0.121	0.128	0.115	0.128	0.11	0.105	0.105
Economic	Internal realization conditions	0.177	0.204	0.239	0.301	0.207	0.197	0.216	0.182	0.204	0.193
	Potential added value	0.431	0.442	0.324	0.163	0.393	0.373	0.293	0.493	0.442	0.471
	Target group willingness	0.104	0.119	0.140	0.177	0.121	0.115	0.127	0.107	0.119	0.113
	Technical realization conditions	0.190	0.219	0.257	0.323	0.222	0.211	0.232	0.195	0.219	0.207
	legal realization conditions	0.098	0.017	0.039	0.035	0.056	0.104	0.132	0.024	0.016	0.015

the already observed domination of the hybrid workload towards the complexity of analysis.

$$\bar{a}_{ij} = \frac{a_{ij}}{\sum_{k=1}^n a_{kj}} \tag{3}$$

$$\bar{a}_{11} = \frac{1}{1 + 5 + 3} \approx 0.111 \tag{4}$$

TABLE V. Sample Calculation of the Weightings

	Complexity	Hybrid Workload	Source Systems	Weight
Complexity	0.111	0.130	0.077	0.106
Hybrid Workload	0.556	0.652	0.692	0.633
Source Systems	0.333	0.217	0.231	0.260

Subsequently, the targeted weightings are determined by averaging the results. The elements are computed as follows:

$$w_i = \frac{\sum_{k=1}^n \bar{a}_{ik}}{n} \tag{5}$$

$$w_1 = \frac{0.111 + 0.130 + 0.077}{3} = 0.106 \tag{6}$$

Calculation of the Consistency

The possibility to detect inconsistent answers helps to ensure the quality of the results. Judgements of persons are often not completely consistent. Especially in complex decision situations the individual judgements do not conform perfectly. The AHP method includes techniques for the detection of contradictory answers. The consistency ratio (CR) calculates the relation between the consistency index (CI) of the matrix and the random index (RI). The equation for the calculation of the consistency index as well as the results of the continued example from the part before can be seen below. The term

λ_{max} represents thereby the maximum eigenvalue of the matrix A [52].

$$CI = \frac{(\lambda_{max} - n)}{n - 1} \tag{7}$$

$$CI = \frac{(3.039 - 3)}{3 - 1} = 0.0195 \tag{8}$$

In the final step, the consistency ratio is calculated. The included term RI is an consistency index for randomly chosen results. It is assumed that the random results are highly inconsistent. The consistency ratio thus compares the quality of the examined example with the quality of arbitrary results. The corresponding values for RI are provided by Saaty in [53], as shown in Table VI.

TABLE VI. Random Indices for different sizes of matrices

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

The unanimous opinion about the consistency is that only answers with a consistency ratio lower or equal 0.1 are considered. In the example shown equation (10), the consistency ratio is about 0.034 and thus can be accepted. Throughout the investigation in overall two data sets had to be revised, due to the exceeding of the consistency ratio.

$$CR = \frac{CI}{RI} \tag{9}$$

$$CR = \frac{0.0195}{0.58} \approx 0.034 \tag{10}$$

The aggregated results of the AHP in Table VII reveal that for the evaluation of the value-creation independent analysis factors the complexity of analysis and the hybrid workload have the main impact. The amount of source systems is in this context only of minor importance. A more notable tendency can be seen between the data volume and the data complexity. The results of this category show, that the significance of the data volume is much higher in comparison to the data complexity.

TABLE VII. WEIGHTINGS OF THE VALUE-CREATION INDEPENDENT FACTORS

Category	Subcategory	Subcategory Weight	Factor	Local Weight
Analysis	Value-Creation independent	0.38	Complexity of analysis	0.42
			Hybrid workload	0.44
	Value-Creation dependent	0.62		
Data	Value-Creation independent	0.55	Data volume	0.81
	Value-Creation dependent	0.45	Data complexity	0.19

C. Evaluation Examples

In the following part, we will apply and interpret the results of the IMIS framework based on three selected use cases. The first chosen example comes from an early adopter of IMIS systems. The analysis of point of sales data in the retail section is one of the first examples in this area. The company participating in our case study is a leading retailer in Germany. For reasons of space and legibility we only show some key attributes of the example. The major goal for the use of an IMIS was the analysis of sales and inventory figures. The use case demands, besides current sales figures, the consideration of fluctuations due to promotions as well as external influences, like the weather conditions. The example is characterized by a high demand regarding the urgency, data volume and the complexity of analysis. The extensive and unpredictable variations of the sales figures require a very fast recognition of anomalies. The calculation includes transactional as well as analytical tasks. Despite, the rare and minor data changes the overall assessment of the data requirements is quite high. This is due to the extremely high volume of data combined with the high weight of this category. The economic evaluation reveal that the most important obstacle concerning the realization of the potential added value is the long implementation duration.

The second example is a real-time reporting case from the insurance sector and is characterized by very high requirements in the analysis as well as in the data area. For this use case, it is necessary that the results are based on up-to-date data and are processed in near real-time. The analyzes are based on large amounts of data directly from the transaction system. The case is characterized by high demands in terms of urgency, degree of detail and the complexity of the analysis. The same accounts for the data factors, showing a extreme high data volume, frequent changes in the data, high effects of the data change and strong changes in the range of variations. From an economic point of view, this case is characterized by a very high added value and a high target group willingness. There are neither internal nor external obstacles that avoid the realization of the results. For these reasons, this example is assessed very high in all categories.

The last example shows very clearly the diverging significance of the influencing factors. The use case comes from a supplier company in the medical field. This company uses IMIS to improve their financial and controlling reports. Despite relatively small changes to the data base, it is important that the data is up-to-date and the results of the analyzes are available very quickly. In comparison to the other use cases, the overall technical requirements are on a lower level. The

same is true for the economical factors. Especially the high legal regimentation stifle/obstruct the economical assessment.

For a better clarity and easier interpretation, we assigned the results of the use case evaluation to a portfolio chart (Figure 4). This chart is comparable to the strategic portfolio matrix of the Boston Consulting Group [54]. The advantage of this representation is the possibility to have a visual indicator for the evaluation of the complex underlying decision problem. The dimensions of the chart are based on the categories of the presented framework. The analysis and data requirements built the axes of the chart. The radius of the data points reflect the economical assessment. The chart is an easy to use tool to indicate promising use cases. The provided example is based on the assessment of 10 use cases, as seen in Table I.

As mentioned at the beginning of this contribution the total utility is calculated using the weighted sum model (11) of the multi-criteria decision making. The calculations of the final metrics is shown for the analysis requirements of use case 1. The weightings are based on the results from the direct ranking and the AHP method (see Sections IV-A and IV-B). The system requirements (x_i) are represented by the results of the use case assessment (as seen in Table I). For the calculation, the assessment have been decoded. The answers are based on a ordinal scale and therefore they can be easily transformed. The sample calculation is shown in (11).

$$U_A = 0.62 \times (0.31 \times 3 + 0.42 \times 5 + 0.27 \times 3) + 0.38 \times (0.42 \times 4 + 0.44 \times 5 + 0.14 \times 4) = 4.21 \quad (11)$$

An interpretation of the calculation only makes sense in comparison to the other results. The visualization indicates that the evaluation of the use cases is quite diverse. The use case *Finance Reporting* for instance, may be characterized by a rather low economical assessment on one side, having medium to low data and analytical requirements on the other side. Although an evaluation of a use case scenario is still subjective to the decision maker's assumptions and weights, the chart provides a tool to either choose, rule out or change

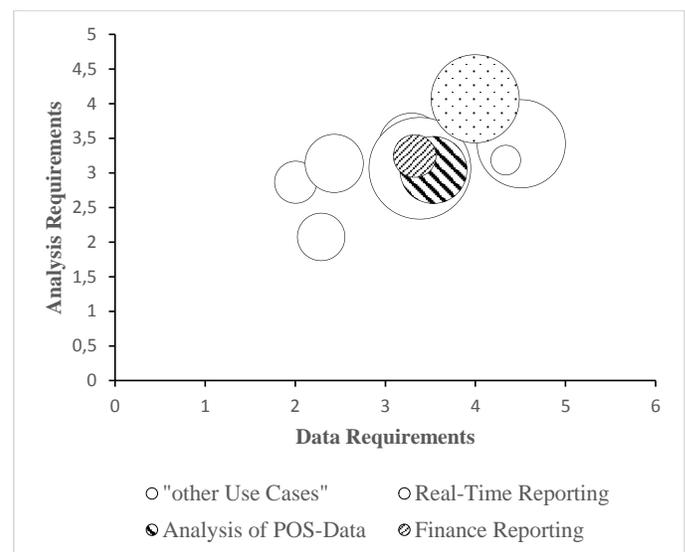


Figure 4. Visualization of the evaluation results

possible use cases. This may also lead to the decision to only use IMIS in parts of the originally planned scenario or to switch to substitute technologies. The process of the application scenario definition, which could be a repetitive process, is also supported by the framework.

V. CONCLUSION

Recent research as well as practical applications of In-memory systems have shown a research gap concerning the structured consideration of IMIS use cases. The aim of this work was the development of a flexible framework for the evaluation of IMIS use cases. Previous IMIS examples have shown a varying importance of the individual influencing factors. To address these variations the decision factors were subdivided into the categories value-creation dependent and value-creation independent. The methods for the determination of the weightings were selected according to these categories and the characteristics of the factors. The weightings of the value creation dependent attributes are determined by the direct ranking method for each investigated use case. For the weightings of the technology oriented, value creation independent factors the analytic hierarchy process was applied. In order to map all factors and their significance to the additive model of the multi-criteria decision making was applied. The presented framework allows to examine existing, as well as exemplary future use cases with regard to the influence factors of In-memory based IT systems. The approach considers both, the system requirements and their corresponding importance. This enables decision-makers to investigate IMIS scenarios for their application potential.

In future work, the framework should be extended to other industries. A broad selection framework is also conceivable that shows reasonable conditions for the use of the In-memory technology. With the aid of the framework, catalogs could be created for suitable and tested application scenarios.

REFERENCES

- [1] S. Ulbricht, M. Opuszko, J. Ruhland, and S. Gehrke, "A use case-oriented framework for the evaluation of in-memory it-systems," in *The Sixth International Conference on Data Analytics*, 2017, pp. 6–11.
- [2] W. H. DeLone and E. R. McLean, "Information systems success: The quest for the dependent variable," *Information systems research*, vol. 3, no. 1, 1992, pp. 60–95.
- [3] J. Peppard and J. Ward, "Beyond strategic information systems: towards an is capability," *The Journal of Strategic Information Systems*, vol. 13, no. 2, 2004, pp. 167–194.
- [4] "Top 10 Hot Big Data Technologies," 2016, URL: <https://www.forbes.com/sites/gilpress/2016/03/14/top-10-hot-big-data-technologies/> [accessed: 2017-07-11].
- [5] "SAP fehlen echte HANA-Business-Cases," 2015, URL: <http://www.cio.de/a/sap-fehlen-echte-hana-business-cases,2940526> [accessed: 2017-07-24].
- [6] "Lack of SAP HANA use cases stifling demand among ASUG members," 2014, URL: <http://diginomica.com/2014/08/08/lack-sap-hana-use-cases-stifling-demand-among-asug-members/> [accessed: 2017-07-24].
- [7] "SAP Business Suite powered by SAP HANA," 2014, URL: <https://www.pac-online.com/download/9757/125462> [accessed: 2017-07-12].
- [8] G. Piller and J. Hagedorn, "Business benefits and application capabilities enabled by in-memory data management." in *IMDM*, 2011, pp. 45–56.
- [9] R. Schütte, "Analyse des einsatzpotenzials von in-memory-technologien in handelsinformationssystemen." in *IMDM*, 2011, pp. 1–12.
- [10] H. Garcia-Molina and K. Salem, "Main memory database systems: An overview," *IEEE Transactions on knowledge and data engineering*, vol. 4, no. 6, 1992, pp. 509–516.
- [11] D. J. DeWitt, R. H. Katz, F. Olken, L. D. Shapiro, M. R. Stonebraker, and D. A. Wood, "Implementation techniques for main memory database systems," in *Proceedings of the 1984 ACM SIGMOD Conference*, vol. 14, no. 2. ACM, 1984, pp. 1–8.
- [12] F. Färber, S. K. Cha, J. Primsch, C. Bornhövd, S. Sigg, and W. Lehner, "Sap hana database: data management for modern business applications," *ACM Sigmod Record*, vol. 40, no. 4, 2012, pp. 45–51.
- [13] D. J. Abadi, P. A. Boncz, and S. Harizopoulos, "Column-oriented database systems," *Proceedings of the VLDB Endowment*, vol. 2, no. 2, 2009, pp. 1664–1665.
- [14] D. Abadi, S. Madden, and M. Ferreira, "Integrating compression and execution in column-oriented database systems," in *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. ACM, 2006, pp. 671–682.
- [15] H. Plattner and A. Zeier, *In-memory data management: technology and applications*. Springer Science & Business Media, 2012.
- [16] H. Plattner, "A common database approach for oltp and olap using an in-memory column database," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*. ACM, 2009, pp. 1–2.
- [17] A. Kemper and T. Neumann, "Hyper: A hybrid oltp&olap main memory database system based on virtual memory snapshots," in *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*. IEEE, 2011, pp. 195–206.
- [18] N. May, A. Böhm, and W. Lehner, "Sap hana—the evolution of an in-memory dbms from pure olap processing towards mixed workloads," *Datenbanksysteme für Business, Technologie und Web (BTW 2017)*, 2017.
- [19] P. Loos, J. Lechtenböcker, G. Vossen, A. Zeier, J. Krüger, J. Müller, W. Lehner, D. Kossmann, B. Fabian, O. Günther et al., "In-memory-datenmanagement in betrieblichen anwendungssystemen," *Wirtschaftsinformatik*, vol. 53, no. 6, 2011, pp. 383–390.
- [20] M. Pezzini, D. Feinberg, N. Rayner, and R. Edjlali, "Hybrid transaction/analytical processing will foster opportunities for dramatic business innovation," *Gartner* (2014, January 28) Available at <https://www.gartner.com/doc/2657815/hybrid-transactionanalytical-processing-foster-opportunities>, 2014.
- [21] F. Özcan, Y. Tian, and P. Tözün, "Hybrid transactional/analytical processing: A survey," in *Proceedings of the 2017 ACM International Conference on Management of Data*. ACM, 2017, pp. 1771–1775.
- [22] J. Krueger, M. Grund, C. Tinnefeld, B. Eckart, A. Zeier, and H. Plattner, "Hauptspeicherdatenbanken für unternehmensanwendungen," *Datenbank-Spektrum*, vol. 10, no. 3, 2010, pp. 143–158.
- [23] P. Loos, S. Strohmeier, G. Piller, and R. Schütte, "Comments on in-memory databases in business information systems," *Business & Information Systems Engineering*, vol. 4, no. 4, 2012, pp. 213–223.
- [24] D. J. Abadi, "Query execution in column-oriented database systems," Ph.D. dissertation, Massachusetts Institute of Technology, 2008.
- [25] J. Krueger, M. Grund, C. Tinnefeld, H. Plattner, A. Zeier, and F. Faerber, "Optimizing write performance for read optimized databases," in *International Conference on Database Systems for Advanced Applications*. Springer, 2010, pp. 291–305.
- [26] T. Winsemann and V. Köppen, "Kriterien für datenpersistenz bei enterprise data warehouse systemen auf in-memory datenbanken." in *Grundlagen von Datenbanken*, 2011, pp. 97–102.
- [27] R. Meyer, V. Banova, A. Danciu, D. Prutscher, and H. Krcmar, "Assessing the suitability of in-memory databases in an enterprise context," in *Enterprise Systems (ES), 2015 International Conference on*. IEEE, 2015, pp. 78–89.
- [28] G. Piller and J. Hagedorn, "In-memory data management im einzelhandel: Einsatzbereiche und nutzenpotentiale," in *Multikonferenz Wirtschaftsinformatik*, 2012.
- [29] —, "Einsatzpotenziale für in-memory data management in betrieblichen anwendungssystemen," *Wirtschaftsinformatik & Management*, vol. 3, no. 5, 2011, pp. 18–25.
- [30] R. Winter, S. Bischoff, and F. Wortmann, "Revolution or evolution?"

- reflections on in-memory appliances from an enterprise information logistics perspective.” in *IMDM*, 2011, pp. 23–34.
- [31] J. vom Brocke, S. Debortoli, and O. Müller, “In-memory database business value,” 360 - The Business Transformation Journal, vol. 3, no. 7, 2013, pp. 16–26.
- [32] J. vom Brocke, “In-memory value creation, or now that we found love, what are we gonna do with it?” *BPTrends*, vol. 10, 2013, pp. 1–8.
- [33] J. vom Brocke, S. Debortoli, O. Müller, and N. Reuter, “How in-memory technology can create business value: insights from the hilti case,” *Communications of the Association for Information Systems*, vol. 34, no. 1, 2014, pp. 151–167.
- [34] R. Bärenfänger, B. Otto, and H. Österle, “Business value of in-memory technology—multiple-case study insights,” *Industrial Management & Data Systems*, vol. 114, no. 9, 2014, pp. 1396–1414.
- [35] C. Cundius and R. Alt, “A process-oriented model to business value—the case of real-time it infrastructures,” in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.
- [36] —, “Real-time or near real-time?-towards a real-time assessment model,” *Thirty Fourth International Conference on Information Systems*, 2013.
- [37] L. Olsson and C. Janiesch, “Real-time business intelligence und action distance: Ein konzeptionelles framework zur auswahl von bi-software.” in *Wirtschaftsinformatik*, 2015, pp. 691–705.
- [38] S. Ulbricht, M. Opuszko, J. Ruhland, and M. Thrum, “Towards an analysis and evaluation framework for in-memory-based use cases,” in *The Twelfth International Multi-Conference on Computing in the Global Information Technology*, 2017, pp. 22–27.
- [39] E. Triantaphyllou, *Multi-criteria decision making methods*. Kluwer Academic Publishers, Dordrecht, 2000.
- [40] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design science in information systems research,” *MIS quarterly*, vol. 28, no. 1, 2004, pp. 75–105.
- [41] A. R. Hevner, “A three cycle view of design science research,” *Scandinavian journal of information systems*, vol. 19, no. 2, 2007, p. 4.
- [42] R. Klein and A. Scholl, *Planung und Entscheidung*. Vahlen, München, 2004.
- [43] D. Von Winterfeldt and W. Edwards, “Decision analysis and behavioral research,” Cambridge, MA (USA) Cambridge Univ. Press, 1993.
- [44] P. Goodwin and G. Wright, *Decision analysis for management judgment*. Wiley, 1998.
- [45] J. R. Doyle, R. H. Green, and P. A. Bottomley, “Judging relative importance: direct rating and point allocation are not equivalent,” *Organizational behavior and human decision processes*, vol. 70, no. 1, 1997, pp. 65–72.
- [46] J. Jia, G. W. Fischer, and J. S. Dyer, “Attribute weighting methods and decision quality in the presence of response error: a simulation study,” *Journal of Behavioral Decision Making*, vol. 11, no. 2, 1998, pp. 85–105.
- [47] W. Edwards and D. von Winterfeldt, “Decision analysis and behavioral research,” Cambridge University Press, vol. 604, 1986, pp. 6–8.
- [48] W. Edwards, R. Miles, and D. von Winterfeldt, *Advances in Decision Analysis: From Foundations to Applications*. Cambridge University Press, 2007.
- [49] T. L. Saaty, “What is the analytic hierarchy process?” in *Mathematical models for decision support*. Springer, 1988, pp. 109–121.
- [50] C.-C. Wei, C.-F. Chien, and M.-J. J. Wang, “An ahp-based approach to erp system selection,” *International journal of production economics*, vol. 96, no. 1, 2005, pp. 47–62.
- [51] M. Godse and S. Mulik, “An approach for selecting software-as-a-service (saas) product,” in *Cloud Computing, 2009. CLOUD’09. IEEE International Conference on*. IEEE, 2009, pp. 155–158.
- [52] M. Kwiesielewicz and E. Van Uden, “Inconsistent and contradictory judgements in pairwise comparison method in the ahp,” *Computers & Operations Research*, vol. 31, no. 5, 2004, pp. 713–719.
- [53] T. L. Saaty, “The analytic hierarchy process: planning, priority setting, resources allocation,” New York: McGraw, vol. 281, 1980.
- [54] D. C. Hambrick, I. C. MacMillan, and D. L. Day, “Strategic attributes and performance in the bcg matrix a pims-based analysis of industrial product businesses,” *Academy of Management Journal*, vol. 25, no. 3, 1982, pp. 510–531.



www.iariajournals.org

International Journal On Advances in Intelligent Systems

🔗 issn: 1942-2679

International Journal On Advances in Internet Technology

🔗 issn: 1942-2652

International Journal On Advances in Life Sciences

🔗 issn: 1942-2660

International Journal On Advances in Networks and Services

🔗 issn: 1942-2644

International Journal On Advances in Security

🔗 issn: 1942-2636

International Journal On Advances in Software

🔗 issn: 1942-2628

International Journal On Advances in Systems and Measurements

🔗 issn: 1942-261x

International Journal On Advances in Telecommunications

🔗 issn: 1942-2601