Michael Freitag
Herbert Kotzab
Nicole Megow *Editors*

Dynamics in Logistics

Twenty-Five Years of Interdisciplinary Logistics Research in Bremen, Germany

Springer

Dynamics in Logistics

Michael Freitag • Herbert Kotzab • Nicole Megow
Editors

Dynamics in Logistics

Twenty-Five Years of Interdisciplinary Logistics Research in Bremen, Germany

Springer

Editors

Michael Freitag
Faculty of Production Engineering
University of Bremen
Bremen, Germany

Herbert Kotzab
Faculty of Business Studies and Economics
University of Bremen
Bremen, Germany

Nicole Megow
Faculty of Mathematics and Computer Science
University of Bremen
Bremen, Germany

ISBN 978-3-030-88661-5 ISBN 978-3-030-88662-2 (eBook)
https://doi.org/10.1007/978-3-030-88662-2

© The Editor(s) (if applicable) and The Author(s) 2021. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book appears on the occasion of the 25th anniversary of interdisciplinary and collaborative research in the Bremen Research Cluster for Dynamics in Logistics (LogDynamics). Hand-picked experts, who are or have been affiliated with LogDynamics, provide retrospect articles, up-to-date research contributions, and future research directions.

History

Twenty-five years ago, in 1996, researchers from different faculties at the University of Bremen (Germany) came together to conduct research in the field of logistics thenceforth jointly in an interdisciplinary way. They established the "Forschungsverbund Logistik" (research association for logistics) to gain a holistic view on logistics systems such as production networks and supply chains, logistic facilities such as warehouses and port terminals, and intra-logistics within companies. Hence, the associated researchers came from economics and business administration, industrial engineering and management, production engineering, computer science, mathematics, and electrical engineering.

In the beginning, the "Forschungsverbund Logistik" focused on the design and management of logistic networks for globally distributed production systems. Activities included application-oriented research projects, a graduate school for the research associates of the involved institutes, chairs, and research groups, a research colloquium, and courses for graduate students at the University of Bremen.

In the 2000s, additional researchers joined the association and developed the concept of autonomously controlled logistics processes. In this context, autonomy means that smart logistic objects receive a certain degree of autonomy and control themselves in a decentralized way, such as cargo in transportation networks or parts and sub-assemblies on the shop floor. This research was conducted within the Collaborative Research Center "Autonomous cooperating logistic processes—a paradigm shift and its limitation" (CRC 637), funded by the German Research

Foundation from 2004 to 2012. The results are reported in more than 100 academic journal papers and more than 300 academic conference contributions.

As autonomous control outperforms conventional planning in situations with high dynamics inside and outside the logistic system, the focus lay on the dynamics of logistics. Accordingly, the association renamed itself in 2005 into "Bremen Research Cluster for Dynamics in Logistics" or in short "LogDynamics."

In addition to the basic research conducted in the CRC 637, LogDynamics established the International Graduate School for Dynamics in Logistics for PhD students from all over the world. Until now, more than 50 international students received their PhD degree as members of this institution.

To transfer research results into practice, LogDynamics also established the LogDynamics Lab. Here, innovative information and communication technologies supporting different logistic processes are tested and optimized toward their application in companies and at logistic service providers.

Autonomous control of logistic processes was one of the foundations for Industry 4.0 and Logistics 4.0. Accordingly, LogDynamics enhanced the developed models, methods, and tools toward cyber-physical production and logistics systems. In the last 10 years, digitalization of logistics was one of the most important research fields of LogDynamics. This included technical developments such as sensors and the subsequent data processing by machine learning algorithms, methodological developments such as data-driven planning, optimization, and control methods, conceptual work on future logistics solutions for the growing e-commerce sector, and new business models.

Recent research interests focus particularly on the trade-off between individualization of logistic services on the one hand and their sustainability on the other hand. Will it be possible to establish end user-tailored home delivery systems and decrease the consumption of resources and the carbon dioxide emission at the same time?

Future research in logistics will be significantly influenced by societal megatrends such as globalization, individualization, connectivity, mobility, urbanization, and neo-ecology. These trends do not appear side by side but rather diametrically, which offers a lot of opportunities for logistics research which deals with trade-offs ever since. Having the advantage of a holistic observation and analysis, LogDynamics is well prepared for the next 25 years of research in the field.

This Book

This book highlights interdisciplinary aspects of logistics research conducted in Bremen. It features articles with empirical, methodological, as well as practice-oriented contributions and addresses modeling, planning, optimization, and control of processes. The articles are clustered in three parts.

The first part addresses "Models and Methods for Planning in Logistics." Schukraft, Teucke, Freitag, and Scholz-Reiter give an overview on the research that BIBA - Bremer Institut für Produktion und Logistik has performed over the past few years in the field of autonomously controlled production and transportation networks. Megow and Schlöter present recent results on mathematical optimization with explorable uncertainty and outline its potential power in the context of decision-making under uncertainty in logistics. In another retrospect article, Becker and Wagner-Kampik revisit selected advances in network modeling and analysis in manufacturing and logistics including also the transition from static to dynamic and stochastic models. Schmand discusses recent developments in the design and analysis of mathematical traffic models.

The focus of the second part is on "Digitalization and Logistics." In the first article, Herzog and Timm demonstrate the power of intelligent software agents as a technological solution for planning, designing, optimizing, and implementing social and learning logistics systems. Hribernik, Franke, and Thoben introduce in their article the semantic mediator and present examples of its successful application to tackle problems of interoperability of heterogeneous information sources in autonomous cooperating logistics processes. Kümpel, Mueller, and Beetz introduce the semantic digital twin for retail as the semantic connection of a scene graph, a model of the environment, with a symbolic knowledge base, allowing for abstract reasoning about its contained facts in various retail applications. Wicaksono, Boroukhian, and Bashyal propose a dynamic demand–response system that aims to engage manufacturing power consumers through price and incentive-based demand–response programs using linked data and machine learning. Knapp, Kessler, and Arlinghaus demonstrate and discuss the influence of cognitive biases on human decision-making in digitalized logistics environments and their effect on logistics performance.

The third part "Fields of Application in Logistics" provides insight into certain application areas of logistics. Hoff-Hoffmeyer-Zlotnik, Teucke, Oelker, and Freitag provide insight to the research within the area of transshipment points and across the downstream supply chain by using specific Industry 4.0 technologies in general and present results of transferring this knowledge to the field of automotive logistics. Jedermann and Lang document 15 years of intelligent container research, which started with the provision of the necessary sensor systems and got applied in food cool chain logistics in order to significantly reduce food losses, especially in the supply chain for bananas. Jungen, Specht, Ovens, and Lemper look at the growing importance of ultra large container vessels for maritime logistics and examine the implications for sea ports as well as environmental issues. Trapp, Luttermann, Rippel, Kotzab, and Freitag examine the sustainability of last mile logistics by modeling individualized home delivery systems for Internet-shopped groceries. Haasis, Du, and Sun present the New Silk Roads and elaborate on the logistical challenges along these new global trade lanes. And finally, Louis and Rügge introduce how the Bremen International Graduate School for Dynamics in Logistics works and show some insights into doctoral research seminars.

The editors would like to thank all authors for their valuable contribution to this collection. Furthermore, the editors would like to thank their LogDynamics colleagues for the fruitful cooperation in the past few years and look forward to the next 25 years of interdisciplinary and collaborative research in the Bremen Research Cluster for Dynamics in Logistics.

Bremen, Germany
Michael Freitag
Herbert Kotzab
Nicole Megow

July 2021

Contents

Part I Models and Methods for Planning in Logistics

Autonomous Control of Logistic Processes: A Retrospective 3
Susanne Schukraft, Michael Teucke, Michael Freitag, and Bernd Scholz-Reiter

Explorable Uncertainty Meets Decision-Making in Logistics 35
Nicole Megow and Jens Schlöter

Complex Networks in Manufacturing and Logistics: A Retrospect 57
Till Becker and Darja Wagner-Kampik

Recent Developments in Mathematical Traffic Models 71
Daniel Schmand

Part II Digitalization and Logistics

Intelligent Agents for Social and Learning Logistics Systems 91
Otthein Herzog and Ingo J. Timm

Semantic Interoperability for Logistics and Beyond 109
Marco Franke, Karl Hribernik, and Klaus-Dieter Thoben

Semantic Digital Twins for Retail Logistics 129
Michaela Kümpel, Christian A. Mueller, and Michael Beetz

A Demand-Response System for Sustainable Manufacturing Using Linked Data and Machine Learning 155
Hendro Wicaksono, Tina Boroukhian, and Atit Bashyal

The Influence of Cognitive Biases in Production Logistics 183
Florian Knapp, Melanie Kessler, and Julia C. Arlinghaus

Part III Fields of Application in Logistics

Automobile Logistics 4.0: Advances Through Digitalization 197
Marit Hoff-Hoffmeyer-Zlotnik, Michael Teucke, Stephan Oelker, and Michael Freitag

15 Years of Intelligent Container Research 227
Reiner Jedermann and Walter Lang

The Rise of Ultra Large Container Vessels: Implications for Seaport Systems and Environmental Considerations 249
Hendrik Jungen, Patrick Specht, Jakob Ovens, and Burkhard Lemper

Modeling Individualized Sustainable Last Mile Logistics 277
Markus Trapp, Sandra Luttermann, Daniel Rippel, Herbert Kotzab, and Michael Freitag

Logistics Challenges Along the New Silk Roads 295
Hans-Dietrich Haasis, Jianhui Du, and Xuejun Sun

Observations and Perceptions on a Doctoral Research Seminar in Engineering and Logistics 309
Dave A. Louis and Ingrid Rügge

Part I

Models and Methods for Planning in Logistics

Autonomous Control of Logistic Processes: A Retrospective

Susanne Schukraft, Michael Teucke, Michael Freitag, and Bernd Scholz-Reiter

Abstract Manufacturing and logistic service companies are increasingly confronted with high dynamics and complexity. Due to its particular suitability for short-term and situation-dependent decision-making, autonomous control can improve planning and control of production and related transportation processes. This chapter gives an overview of the research that the BIBA—Bremer Institut für Produktion und Logistik GmbH has performed over the past years in the field of autonomously controlled production and transportation networks. The chapter focuses on the modeling approaches and the autonomous control methods that have been developed. These methods have been evaluated using both theoretical and real-world scenarios. The results show the applicability and suitability of autonomous control in complex and dynamic production and transportation environments. In addition, influences on the methods' performance and the integration of autonomous control into conventional planning and control systems are discussed. Finally, the chapter looks at the significance of autonomous control in the context of Industry 4.0 and shows the relations between both concepts.

S. Schukraft (✉) · M. Teucke
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany
e-mail: skf@biba.uni-bremen.de; tck@biba.uni-bremen.de

M. Freitag
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

Faculty of Production Engineering, University of Bremen, Bremen, Germany
e-mail: fre@biba.uni-bremen.de

B. Scholz-Reiter
Faculty of Production Engineering, University of Bremen, Bremen, Germany
e-mail: bsr@biba.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_1

1 Introduction

Over the last decades, manufacturing companies have been confronted with increasing dynamics and complexity (Alkan et al. 2018; Duffie et al. 2017). Among others, companies were challenged with volatile market conditions, decreasing product life cycles, and an increasing product variety (Nyhuis and Wiendahl 2009). The main objective of manufacturing companies is to meet customer requirements regarding product quality and agreed delivery dates (Hon 2005). Therefore, an essential issue for production planning and control is the ability to react flexibly to changing conditions and requirements, e.g., changes in customer demands regarding the product, its volume, or its delivery date. Furthermore, manufacturing companies depend on reliable transportation processes. Thus, to fulfill customer demands, the need for a high degree of flexibility also applies to the related transportation processes for supplying raw materials and delivering finished products to the customer.

Up to now, production planning and control is mainly based on centralized and hierarchical planning approaches. Generally, these approaches calculate production schedules in advance, assuming static production environments. Strategies to cope with changing conditions include the partial or complete rescheduling of the production schedule in the case of occurring demand changes or the provision of robust schedules in advance. However, whereas recurring rescheduling activities may lead to a high planning nervousness, working with robust schedules may be inefficient when the flexibility provided is not utilized (Ouelhadj and Petrovic 2009).

An opposite approach to overcome the limitations of these conventional planning and control approaches is the development of autonomously controlled production systems. Generally, autonomous control is based on decentralized decision-making based on the current system state (Freitag et al. 2004). Thus, autonomous control methods are enabled to flexibly react to changing conditions and achieve a high level of target compliance even in highly complex and dynamic environments.

The BIBA—Bremer Institut für Produktion und Logistik GmbH at the University of Bremen as a member of the Bremen Research Cluster for Dynamics in Logistics (University of Bremen 2021)—has a long history of research in the field of autonomous control of production and logistics, including both manufacturing systems and transportation networks. This paper gives an overview of this research, some of which has also been carried out in collaboration with other research and industrial partners. The presented work mainly bases on the results of the Collaborative Research Center 637 “Autonomous Cooperating Logistic Processes: A paradigm Shift and its Limitations” (CRC 637). The CRC 637 has been carried out at the University of Bremen from 2004 to 2012 with several research departments at the University of Bremen and also the Jacobs University of Bremen. In addition, the research was continued in several follow-up projects that are also considered in the given overview.

This chapter is structured as follows. In Sect. 2, autonomous control and its main characteristics, as well as basic assumptions regarding the performance of autonomous control in complex environments, are introduced. In addition, the section discusses the relations between autonomous control and the Industry 4.0 paradigm. An overview of the reference scenarios and modeling approaches for the development of autonomous control methods is given in Sect. 3. Section 4 introduces autonomous control methods developed for either production or transportation scheduling. Section 5 presents the results of various simulation studies mainly based on theoretical reference scenarios. The transfer of autonomous control into real-world applications is described in Sect. 6. A conclusion is given in Sect. 7.

2 The Concept of Autonomous Control

Autonomous control is characterized by processes of decentralized decision-making within a heterarchical organization structure. Thus, the decision-making authority is delegated to individual system elements allowed and enabled to make and execute decisions independently on their own (Freitag et al. 2004). In doing so, the logistic objects require the ability to interact with each other to gather relevant decision-making information. The characteristics of heterarchy suggest a high degree of independence between individual system elements and the central coordination unit. Furthermore, autonomously controlled systems are characterized by non-determinism, which means that the system behavior cannot be predicted over a longer period of time, even if all system states are known (Windt and Hülsmann 2007).

The autonomy of logistic objects necessary for decentralized decision-making is realized through information and communication technologies (ICT). The automatic identification of logistic objects requires technologies like barcodes or radiofrequency identification (RFID). Terrestrial technologies can be used for the automatic localization of logistic objects, either satellite-based global navigation satellite systems (GNSS) or long-range, short-range, and indoor terrestrial technologies. Sensor units allow condition monitoring of the logistic objects and their environment. The information processing can be supported by multi-agent systems (MAS) or machine learning algorithms. Technologies like past and recent mobile phone standards or Long Range Wide Area Networks (LoRa) can be used for the information exchange between logistic objects (Böse and Windt 2007b). Altogether, the significant developments in ICT over the last decades that culminate in the introduction of Industry 4.0 can serve as an enabler for implementing autonomously controlled production and logistics systems in practice.

ICT such as RFID or Wi-Fi networks were emerging when the research of the BIBA on developing autonomous control started in 2004 but still not uniformly introduced into the economy. However, it was anticipated that these technologies would gradually become more widely available and used in industry. From 2011 onwards, the concept of Industry 4.0 (Kagermann et al. 2013) emerged and soon established itself internationally as a collective term for the ongoing digitalization and smart automation in industrial production and the integration of the underlying technologies. From this moment on, some of the technological developments that were integral to key concepts of autonomous control were subsumed under the generic term Industry 4.0.

Considering the principles of both concepts, it can be stated that autonomous control and Industry 4.0 are closely related to each other. The ability of an autonomous object to process information and render decisions on its own necessitates intelligence embedded into this object in the form of data processor units and sensors that provide data on the object's state and its environment. Also, the interaction between several such autonomous objects requires their ability to exchange data, which presumes them being linked to each other via communication networks. These concepts are closely mirrored by two key principles integral to Industry 4.0, namely decentralized decisions and interconnectivity (Hermann et al. 2016). Decentralized decisions in Industry 4.0 are based on the ability of the so-called cyber-physical systems (CPS). CPS are characterized by the interconnection of information technology and software components with mechanical and electronic parts that communicate via a data infrastructure formed by wired or wireless communication networks (Broy 2010). One example mentioned for decentralized decision-making in Industry 4.0 was products routing themselves through a production system (Kagermann et al. 2013), which was intensively studied in the context of autonomous control (see Sect. 4.2.1). Interconnectivity denotes, in the context of Industry 4.0, the ability of machines, devices, sensors, and people to connect and communicate with each other via the Internet of Things (IoT). The concept of IoT describes the technologies for a global ICT infrastructure that enables the connection of physical and virtual objects and the globally unique identification and addressing of these objects, allowing them to interact and collaborate (Bouhai and Grayson 2017; Chaouchi 2013). In sum, it can be stated that Industry 4.0 provides the technological basis and leads to the widespread use of technical systems that enable autonomously controlled logistics. Autonomous control provides methods and algorithms that can lead to an increase in efficiency in production and logistics, particularly in complex and dynamic environments.

Autonomous control methods take the current system state into account for decision-making. Consequently, these methods are considered especially promising if changing conditions require a flexible reaction of planning and control, e.g., caused by workstation breakdowns or short-term rush orders. However, conventional planning methods are able to create optimal production schedules in advance (i.e., before production orders are released to the shop floor). These methods are often computationally intensive but can outperform autonomous control in static production environments. Furthermore, entirely autonomously controlled production systems also have the risk of a chaotic and unpredictable system's behavior (Windt and Hülsmann 2007). Consequently, potential users of autonomous control will require assistance to assess in which cases and to which extent the application of autonomous control is promising and in which cases conventional planning methods still outperform autonomous control. To address these questions, the following hypotheses were formulated regarding the performance of autonomous control. These hypotheses will be revisited below when discussing the methods' performance and comparing them to conventional planning methods.

1. Autonomous control methods are able to cope efficiently with complex and dynamically changing production environments (Freitag et al. 2004).
2. In static production environments, conventional planning and control methods will reach an equal or even better performance than autonomous control. Contrarily, autonomous control will outperform conventional planning and control in environments of high complexity (Philipp et al. 2006).
3. Besides hypothesis (2), the performance of autonomous control methods depends on the present degree of autonomy and the complexity of the applied scenario (Böse and Windt 2007b; Windt et al. 2008). A more detailed description of the assumed interdependencies is given by Philipp et al. (2006).
4. A possibility to combine the advantages of conventional planning and autonomous control—the high planning accuracy of conventional planning in static and the high efficiency of autonomous control in dynamic environments—is the combined application of both approaches (Schukraft et al. 2016a).

3 Reference Scenarios for Autonomous Logistic Processes

This section introduces reference scenarios for both production and transportation scheduling. These theoretical scenarios provide a standardized environment that allows evaluating and comparing the different autonomous control methods' performance. In addition, the section describes the simulation and evaluation of autonomous logistic processes.

3.1 Description of the Reference Scenarios

The reference scenarios have been defined in a way that they can easily be adjusted to analyze the influence of various impact factors (e.g., different system sizes or variation of processing times) on the autonomous control methods' performance. Besides, the modeling approaches for the implementation of the autonomous control methods are introduced. Also, this section describes approaches for the measurement of the autonomous control methods' performance.

3.1.1 Shop-Floor Scenario

The "NxM shop-floor scenario" is based on a flexible flow shop system. It consists of M stages, each with N parallel workstations and an input buffer in front of each workstation. Production orders have to be processed on each stage on one of the

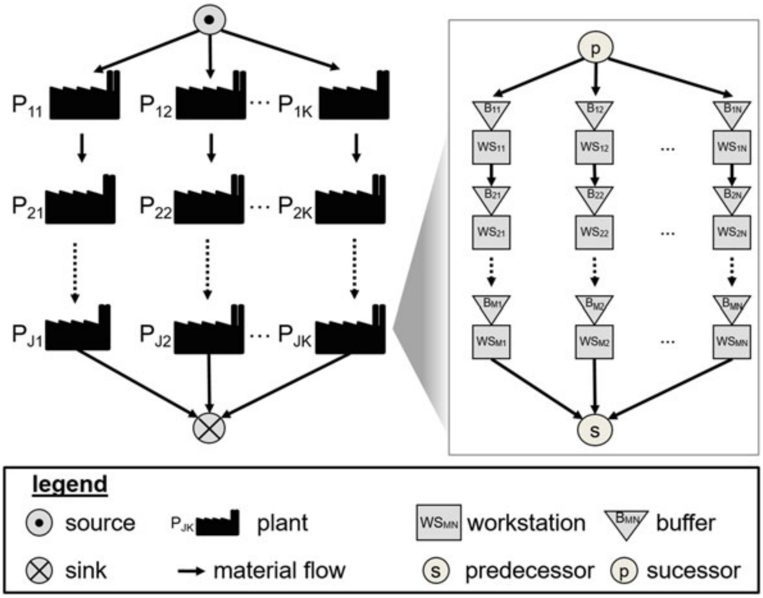


Fig. 1 Production network scenario (based on Scholz-Reiter et al. (2009d))

parallel workstations. Mainly, the scenario was applied as a 3 × 3 scenario, i.e., three stages and three parallel workstations per stage (Scholz-Reiter et al. 2005a, b). However, the model is scalable to any desired problem size, e.g., Scholz-Reiter et al. (2010a) applied the scenario with a maximum of 10 × 10 workstations.

Based on the described shop-floor scenario, a generic "J × K production network scenario" was developed, allowing the integration of external transportation between different production plants (see Fig. 1). Analogous to the shop-floor scenario, the network scenario consists of J × K production plants, i.e., each product has to be processed in one production plant on each production step. Each production plant consists of a shop-floor model with N × M workstations. A transport system connects the production plants. The scenario allows the definition of symmetric and asymmetric network configurations (Scholz-Reiter et al. 2009d).

Besides the scalable model size, both the shop-floor and the network reference scenarios offer various possibilities to consider different influencing factors on the methods' performance. Among others, the reference scenarios allow the variation of processing and set-up times, the consideration of different job types, fluctuation of order entrance, or disturbances like workstation breakdowns.

3.1.2 Transportation Scenario

The transportation scenario covers general cargo transport (package freight) through a dynamic multi-modal transportation network. For the representation of the route network, directed graphs were used (Wenning et al. 2007a). The transportation network consists of a discrete number of geographically distributed vertices (nodes) connected by edges. A vertex (node) represents a fixed location in the transportation network, where several edges (roads) intersect, or trans-shipment is carried out or both. Vertices can be either active or passive. Active vertices include trans-shipment facilities, where transport units are loaded and unloaded. Passive vertices have no trans-shipment facilities. At vertices, sources are located, i.e., departure points of packages that start their journey through the transportation network. Sinks, or destinations, where packages finish their journey through the transportation network, are located at vertices as well. Edges are connections between neighboring vertices, representing, e.g., roads. They are considered to be directed and of a fixed length. A route is a connection between two different vertices that may include one or several edges. Often several different routes are possible to connect two nodes.

Packages represent pieces of goods or cargo that need to be transported from a given source to a given sink in the transportation network. A package is indivisible, which means that it cannot be broken down into smaller parts. Packages cannot move through the transportation network themselves and thus need transport units to move them between nodes of the network. A package can be loaded onto, or unloaded from, a transport unit at active vertices, allowing the exchange of transport units. It can also wait at the vertex for transport in the future. A transport order contains all information needed for carrying out the transport of a package or a group of packages. Transport units are generalized representations of transport units, e.g., trucks or vans. Transport units move between different vertices at a certain speed using the connecting edges. Different types of transport units can be restricted to different subsets of the vertices. Transport units can each take a number of cargo packages with them, which is limited by their specific capacity. At active vertices with trans-shipment capacity, some of the cargo packages can be exchanged.

The transport scenario may include dynamic and stochastic variations: New packages with transport orders are generated at run time, while packages delivered to their destination disappear. Vertices may newly appear or disappear, while edges may be temporarily unavailable for use. The speed of transport units traveling between vertices may be stochastically distributed.

Several instances of this general scenario were implemented, including a small scenario of four vertices and a so-called Germany scenario, where the location of the nodes is based on a network of 18 cities in Germany, while the edges between the vertices represented highway connections between those cities (see Fig. 2). Also, larger transportation networks of up to 87 vertices have been derived from a map of Central Europe (Wenning et al. 2007b; Wenning 2010).

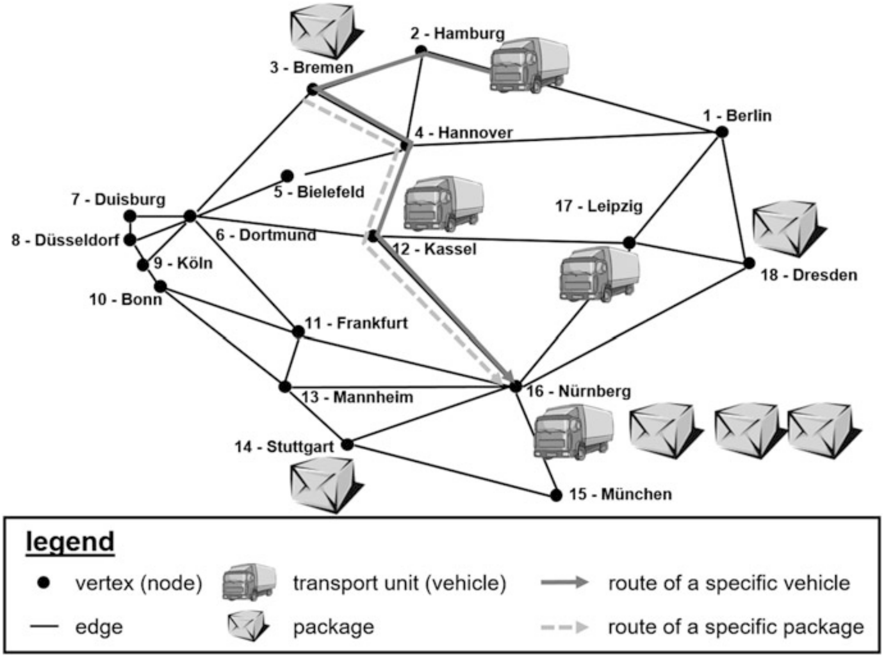


Fig. 2 Transportation scenario (based on Wenning et al. (2007b))

3.2 *Modeling Approaches*

For the modeling of autonomous control methods, mainly discrete-event simulation using the software tool eM-Plant (today known as Tecnomatix Plant Simulation) has been applied. Both discrete-event simulation and the software Plant Simulation are well established and often used for the simulation of material flow processes in manufacturing and logistics. The discrete-event modeling approach was also compared to continuous modeling using the also established System Dynamics software tool VENSIM. The comparison showed that discrete-event simulation allows a detailed description of the shop floor processes but requires high programming effort to implement the autonomous control strategies. These strategies can be implemented in continuous models rather quickly, but the shop floor processes are described only on a rather high aggregation level (Scholz-Reiter et al. 2005b).

A combination of both discrete-event and continuous approaches has been applied to analyze the stability of autonomously controlled production systems. Roughly speaking, a production system is stable if its state remains bounded over time. Therefore, a mathematical investigation using differential equations was first applied to calculate the parameters for which stability of the production system can be guaranteed. Second, the results of the mathematical stability analyses are used as a starting point for discrete-event simulation. With an enlarged set of parameters, the discrete-event simulation can be used to refine the calculated stability results. This combined approach takes advantage of both continuous and discrete-event modeling approaches. The main advantage lies in the high time efficiency. Since the discrete-event model uses the results of the mathematical approach as input, the detailed stability parameters can be determined in less time compared to a pure trial-and-error simulation study (Scholz-Reiter et al. 2011d).

In addition to using established software tools, the scientists within the CRC 637 also developed modeling tools that support the modeling of autonomously controlled processes in particular. A multi-agent simulation platform, the Platform for Simulations with Multiple Agents (PlaSMA), allows simulating mutual negotiations of software agents, which can represent autonomous objects (Warden et al. 2010). Also, a simulation environment for the modeling of autonomously controlled transportation processes was developed. The simulation environment uses a C++ class library for the discrete-event simulation of communication networks (Becker et al. 2006). With the Autonomous Logistics Engineering Methodology (ALEM), a framework for the modeling of autonomous logistic processes and their infrastructure was developed (Scholz-Reiter et al. 2007b).

3.3 Evaluation of Autonomously Controlled Systems

The profound evaluation of the performance of autonomous control methods requires a target system that contains performance measurements for the basic categories of logistic objectives: due date adherence, throughput time, work in process, and utilization (Nyhuis and Wiendahl 2009). Therefore, Philipp et al. (2007) introduced a vectorial approach that considers these basic objectives with weighting factors to adjust the influence of each objective onto the total logistics target achievement. Grundstein et al. (2015b) enhanced the vectorial approach to additionally consider measurements of planning adherence.

The analysis of the assumed influences on the autonomous control methods' performance (cf. Sect. 2) requires measuring the scenarios' autonomy and complexity. In this context, Böse (2012) introduced a morphological pattern to specify a production system's level of autonomy. The measurement contains criteria within the categories "decision-making," "information processing," and "decision execution" and, thus, covers the fundamental characteristics of autonomous control. Grundstein et al. (2015b) expanded the morphological pattern to consider both the planning and the control level. In detail, the expansion focused on criteria to describe coupling strategies used for a combined application of conventional planning and autonomous control (see Sect. 4.3).

The measurement of the production systems' complexity considers various dimensions of complexity (Windt et al. 2008). A complexity cube describes the system's complexity within the dimensions "organizational complexity," "time-related complexity," and "systemic complexity." Each area of the complexity cube can be described by a complexity vector, resulting in a complexity value for the considered production system.

The described approaches for the performance measurement and the specification of the levels of autonomy and complexity have been integrated into a so-called three-component-evaluation system (Philipp 2014; Windt et al. 2010c). The authors also stated detailed assumptions regarding the interdependencies between the dimensions of the evaluation system.

4 Methods for Autonomous Control

This section introduces the methods developed for autonomously controlled production systems and transportation networks. This includes, first, a method for the modeling of logistic processes and their required infrastructure components. In addition, methods for decision-making for both production and transportation and the coupling of autonomous control with conventional planning methods are introduced.

4.1 Engineering Methodology for Autonomous Logistic Processes

For the realization of autonomously controlled systems, the logistic objects and their required infrastructure components have to be modeled. These components depend on the selected control strategy and the architecture of the control system. In this context, the Autonomous Logistics Engineering Methodology (ALEM) offers a procedure model for configuring autonomous logistic processes and their infrastructure. The process mapping is based on a view concept that distinguishes between five views of modeling (Scholz-Reiter et al. 2007b). The “structure view” contains the representation of the logistic objects as well as their interrelationships. The required knowledge of the logistic objects for decision-making is described in the “knowledge view.” The “capability view” concretizes the capabilities of the individual logistic objects. The “process view” is used to model the logical and temporal sequence of activities and states of the logistic objects. Finally, the “communication view” focuses on the content and temporal sequence of the information exchange between the logistic objects (Scholz-Reiter et al. 2007b).

The notation within the described views is based on the Unified Modeling Language (UML). UML includes various diagram types, which can be divided into structure diagrams and behavior diagrams. The static structure is represented in ALEM within the structure, the knowledge, and the capability view. Within these views, class diagrams and object diagrams are used. The description of dynamic processes in ALEM comprises the process and the communication view.

In the process view, activity diagrams are used to describe the process flows. The communication view uses sequence diagrams to specify the temporal and logical sequence of communication between different objects (Scholz-Reiter et al. 2007b).

In addition to the described view concept and the UML-based notation, ALEM provides a procedure model to support the user during modeling. In summary, ALEM offers a set of methods and tools that have been combined within a software tool to support the mapping of autonomous logistic processes (Sowade et al. 2012).

4.2 Autonomous Control Methods

This section introduces different methods for autonomous decision-making in production systems and logistics networks. An overview of the developed methods is given in Table 1. Generally, the methods can be categorized according to the reference scenarios (cf. Sect. 3.1) in methods for the scheduling of shop-floor, transportation, and network processes (introduced in Sect. 4.2.2).

Besides the application area, the methods also differ in their characteristics. A common differentiation criterion is the information basis used for decision-making. Whereas some methods base their decision on future data (e.g., estimated processing times), others decide based on past data (e.g., recorded processing times of preceding production orders). Also, it can be differentiated between local information and information discovery methods depending on the number of planning steps considered for decision-making. A comprehensive overview of possible criteria to classify autonomous control methods is given by Windt et al. (2010a).

4.2.1 Production Scheduling

The queue length estimator (QLE) chooses the workstation based on the estimated waiting time, taking into account the planned processing time of each production order in the input buffer (Scholz-Reiter et al. 2008a, b). Originally, the QLE method for workstation assignment is combined with a first in first out (FIFO) control strategy, which means that already waiting production orders in the buffer are sorted according to their arrival time. A variation called due date method (DUE) is the combination of QLE with sorting the waiting orders according to their due date (Scholz-Reiter et al. 2009c). The bio-inspired pheromone method (PHE) mimics the behavior of ants when foraging. Ants find the shortest way to food by leaving pheromone trails in the environment. On the shortest and thus most frequented path, the pheromone trail is renewed most often and is most pronounced after some time. The ants follow the trail with the highest pheromone concentration. The described principle is transferred to production control as follows: Every part leaves data about the duration of the waiting and processing time when leaving a workstation. Subsequent parts make their decision based on the data left by similar products.

Table 1 Overview on developed autonomous control methods

Autonomous control method		Description	Scenario		
			Shop-floor	Transportation	Network
QLE	Queue length estimator	The method decides on the next workstation based on the estimated waiting time.	X		X
PHE	Pheromone based approach	The method's decision on the next workstation is based on a pheromone trail left by similar products.	X		X
BEE	Bee foraging	The bio-inspired method decides over the next workstation based on the workstations' attractiveness calculated using a so-called value-added calculation.	X		
CHE	Bacterial chemotaxis	The bio-inspired method is based on the concept of chemotaxis and aims to reduce the cycle time.	X		
APC	Autonomous production control method	The method integrates order release, sequencing of production orders, and capacity control and aims at due date adherence and balancing work-in-process.	X		
TUC	Transport unit-centered autonomous control method	Transport units select routes for round trips based on information on waiting packages received by the vertices of a transportation network. Packages are passive objects. The method is based on ad-hoc route-finding mechanisms in communication networks and aims for due date adherence.		X	
DLRP	Distributed logistics routing protocol	The method combines the selection of routes by objects of different types (either routing of transport units and cargo packages through a transportation network or routing of orders through a shop-floor) and can be coupled with different aims (objectives), including, e.g., short transport times, due date adherence, high utilization of transport units	X	X	

The frequency of updating or the number of data records taken into account mimics the evaporation of pheromones in ants (Armbruster et al. 2006; Scholz-Reiter et al. 2008b).

Both the QLE and the PHE methods have been applied to shop-floor scenarios with and without set-up times. Since the order sequence of waiting parts is not communicated by the pheromones, for the PHE method, a correction was used to integrate set-up times into the decision-making. Also, a weighted combination of the QLE and the PHE method, including the mentioned correction term, was developed (Scholz-Reiter et al. 2007a). The QLE and the PHE method are originally developed for production scheduling. However, both methods have been adapted for the assignment of production orders to plants in the production network scenario. The QLE method for plant assignment (QLE_n) decides over the next plant, based on information about the duration of transport, waiting times, and processing times in the subsequent plants. Analogously, the PHE method (PHE_n) for plant assignment decides based on past data about the necessary time to pass the transport system and the corresponding plant (Scholz-Reiter et al. 2009d).

The bee foraging approach (BEE) mimics the behavior of bees in the form of a cost-based approach. For this purpose, the attractiveness of a station for an order is calculated using a “value-added” calculation. The expected lead time and the cost rate of the station are included in this calculation (Scholz-Reiter et al. 2008b). The method of bacterial chemotaxis (CHE) is a bio-inspired method based on the concept of chemotaxis. The method aims to reduce the cycle time (Scholz-Reiter et al. 2010c).

So far, the presented autonomous control methods for shop-floor scheduling focus on workstation assignment combined with certain priority rules for order sequencing (mostly FIFO). In the simulation studies presented below, orders are mostly released to specific times that are set to a sine function. Generally, the methods could be combined with any order release method, but the influence of different order release methods on the autonomous control methods’ performance was not investigated. The due date-oriented autonomous production control method (APC) closes this gap by integrating order release, sequencing of production orders, and capacity control, and thus, all production control tasks. The APC method aims to meet due dates and balancing the work-in-process on a defined level. The method requires basic planning data, e.g., processing start and completion times. Production orders are released if the planned release date is reached or if the range on a workstation on the first stage falls below a defined limit. Workstation assignment is based on due dates and throughput time to balance the workload between parallel workstations. Capacity control is triggered if the defined range limits are violated. The sequencing of production orders takes into account due dates and set-up times. The method’s decisions are based on both future and past information (Grundstein et al. 2017; Grundstein 2017).

Another autonomous control method for production scheduling is the distributed logistics routing protocol for production ($DLRP_p$). The DLRP was originally developed for the autonomous routing of both vehicles and packages in transportation networks (see Sect. 4.2.2 below). However, the basic principles are independent of the original application scope and also hold advantages for production scheduling. Consequently, the DLRP was transferred to production scheduling with the basic idea that production orders should route themselves through a production system that consists of several workstations used for different consecutive operations (Rekersbrink et al. 2010).

4.2.2 Transportation Scheduling

The developed autonomous control methods for transportation scheduling are based on source routing mechanisms used for mobile ad-hoc networks like Wireless Local Area Network (W-LAN or Wi-Fi). Route discovery mechanisms for ad-hoc networks can be divided into proactive, reactive, and hybrid procedures (Perkins 2001). In proactive procedures, each node maintains a table with routes to the other nodes in the network. The nodes exchange appropriate messages about their current state to keep these tables updated. Reactive methods, like dynamic source routing (DSR), determine routes only when needed. The node that wants to communicate sends a route request message to its neighboring nodes. The neighboring nodes forward this route request until the target node is found. The target node responds with a route reply message to the initiator of the routing process, which is propagated back through the network. Hybrid procedures combine mechanisms of proactive and reactive ones. For example, a neighborhood zone can be defined, where most communication is expected to take place and routing is proactive, whereas routes to nodes outside the zone are determined reactively.

Wenning et al. (2005, 2006, 2007b) developed a pure transport unit-centered autonomous control method. Based on the method, each individual vehicle can make its autonomous decisions based on its specific objectives, choosing both its route and the packages to be taken. Vehicles planning their next movements use route request mechanisms to determine a round trip route to the destination vertices of the packages present at a source vertex. The packages are passive objects without intelligence. A genetic algorithm is used to optimize a vehicle target function, which includes transport costs and rewards for the transportation of packages close to their ordered delivery dates.

The distributed logistics routing protocol for transports (DLRP_t) is a combined vehicle-centric and unit-goods-centric approach, which considers both transport units and packages as autonomous, intelligent objects and determines interdependent routes for both types of objects. It combines hybrid routing mechanisms for ad-hoc networks (e.g., DSR), which do not need to know the network topology, with a reactive planning component for the rescheduling of logistic objects (Rekersbrink et al. 2009; Scholz-Reiter et al. 2006a, 2009a).

Both transport units and packages present at a vertex register their destinations at that vertex. The vertex sends route requests for both types of objects to determine possible routes to their respective destinations. The route request and reply mechanism assume that a vertex only has information about itself and its direct neighboring nodes. After receiving several route proposals, the objects can decide on a route and register it with the involved vertices. Different route decision functions can be realized with the DLRP_t. The information required for the decisions must be specified precisely in the RouteRequest and RouteReply messages so that each object can find the necessary information (Scholz-Reiter et al. 2006a). A method based on fuzzy logic is used to evaluate the different proposed routes (Rekersbrink et al. 2007).

Wenning et al. (2010, 2011) generalized the DLRP_t to a more abstract autonomous control method that considers weighted combinations of different decision criteria, such as monetary costs caused by individual logistic objects, ecological costs in the form of carbon dioxide emissions, and risks (e.g., the risk of delays).

In many real-life transport processes, organizational information constraints often restrict the sharing of necessary data. To take these into consideration, Scholz-Reiter et al. (2010b) defined different actors (suppliers and transporters) and their tasks in the logistics network, including the required minimum information. To improve the scalability of information processing in large networks, various mechanisms to limit the propagation of transport requests can be added to the DLRP_t. These include limitation of the search depth, intermediary assessment of partial routes for segments of the total route, re-using data from previous request propagations (Wenning et al. 2009; Wenning 2010), or mechanisms for clustering and regional partitioning of transportation networks (Singh et al. 2008).

4.3 Coupling of Conventional Planning and Autonomous Control

The concept of coupling aims at combining the high efficiency of autonomous control in dynamic production environments with the high planning accuracy of conventional planning methods in static production environments. Thus, conventional planning and autonomous control methods must be enabled to exchange decision-relevant information. This information exchange allows the consideration and execution of necessary adaptions due to occurring dynamic influences and resulting deviations of predefined planning results. Coupling strategies define the extent and the duration of coupling. The extent specifies to which degree autonomous control refers to central planning parameters for decision-making (e.g., predefined workstation assignments or sequences of production orders). The duration defines if coupling is applied permanently or depending on the system state. In the latter case, production orders usually follow the predefined schedule, and autonomous control is only applied if deviations of the schedule are expected. The described coupling strategies have been applied to couple different conventional planning heuristics with several autonomous control methods for production scheduling (Schukraft et al. 2016a).

5 Evaluation of Autonomous Control

The autonomous control methods presented above have been evaluated in various simulation studies using either discrete-event or continuous modeling approaches for the shop-floor scenario. For the transportation scenario, also multi-agent system simulations were used. The simulation results presented in the following are mainly based on the introduced reference scenarios for production systems and transportation networks (cf. Sect. 3.1). For production scheduling, the shop-floor scenario was mainly applied as a 3 × 3-model, but also larger problem sizes up to 10 × 10 workstations were used. Most studies considered various job types, heterogeneous processing times, and sequence-dependent set-up times. Mostly, order entry was modeled using a certain arrival rate set to a sine function. This also allowed a variation of the dynamic influence by adjusting period and amplitude of the sine function. The simulation studies performed to assess the autonomously controlled transportation methods used transportation networks with 18, 40, or 87 vertices. Different criteria were used as performance indicators. These included aggregate transportation times and distances in relation to the shortest paths, delivery time adherence of packages, utilization of transport units, and numbers of queued packages.

5.1 Performance of Autonomous Control and Influence of Complexity

The performance of autonomous control methods has been compared in different simulation studies. Several studies focused on the QLE and the PHE algorithm, which differ in their information basis used for decision-making (future vs. past data). Other studies compared the QLE method with the DLRP$_p$ to compare local information and information discovery methods.

The QLE outperforms the PHE method in high dynamic environments in both the shop-floor scenario (Scholz-Reiter et al. 2006b, 2008b) and the network scenario (Scholz-Reiter et al. 2011b). The PHE performs best in situations with fewer dynamics in the arrival rate. This could be explained since the PHE method's decision-making is based on past data, and thus, the method needs time to react to changing conditions in dynamic environments. Also, a weighted combination of both methods for decision-making leads to a promising performance (Scholz-Reiter et al. 2007a; Scholz-Reiter and Jagalski 2008).

The application of the QLE and the DLRP$_p$ in a shop-floor scenario with up to 10 × 10 lines and workstations showed that both methods reach a comparable performance with slightly better values for the QLE method. It was assumed that the DLRP is designed for more complex environments, and obviously, the simpler decisions of the QLE are more efficient for the used shop-floor scenario (Scholz-Reiter et al. 2010a, 2011c).

Some studies also analyzed the impact of different complexity levels on the methods' performance. Scholz-Reiter et al. (2006c) applied the QLE and the PHE method in a shop-floor scenario. Several complexity levels were considered by using different network sizes and a different number of job types. The QLE method is not affected by a larger system size since the throughput time was always close to the minimal throughput time (calculated with minimal processing time for each production stage). Contrarily, the performance of the PHE algorithm gets worse with increasing system size. For an increasing number of product types, the QLE also outperforms the PHE method. However, the performance of the PHE gets better with a rising number of product types.

Grundstein et al. (2015a) analyzed the impact of different order release methods on the performance of autonomous control. The study applied the order release methods immediate order release (IMR), due date oriented (DATE), and constant work-in-process (CONWIP) in combination with the autonomous control methods QLE, QLE with sorting the production according to the due date (DUE), and PHE. The method combinations were applied in a shop-floor scenario with different levels of dynamics for order arrival. The results show no dominant order release method. Instead, the performance depends on the dynamic level for order arrival and the pursued logistic objective.

The APC method was applied in a shop-floor scenario. The complexity was varied by using different types and intensities of dynamic influences. The method's performance was, among others, compared with the QLE method and conventional planning for the task of workstation assignment. The results show that the performance for all method combinations decreases with an increase of dynamic influences. The application of APC for all control tasks outperforms other method combinations for customer-oriented target profiles with a high weighting of due date-oriented objectives. For other profiles that favor low work-in-process, a combination with the QLE method for workstation assignment reaches higher values. This is explained by the fact that the APC method consciously aims at high due date adherence at the cost of high work-in-process (Grundstein 2017). The results also confirmed the findings of a previous simulation study that used a similar scenario but fewer method combinations and dynamic influences (Grundstein et al. 2015c).

The influence of autonomy and complexity on the performance of autonomous control methods has also been analyzed by using real production data from a power plant supplier. The production system consists of several work systems, each with a different number of workstations. The levels of complexity and autonomy have been defined using the three-component evaluation system described in Sect. 3.3. Different studies focused on different aspects to define the levels of complexity and autonomy. Philipp (2014) and Windt et al. (2010c) addressed complexity by varying the number and diversity of workstations, orders, material flow connections, and manufacturing processes. The level of autonomy varied depending on the applied planning and control logic. Scholz-Reiter et al. (2009c) kept the shop-floor scenario constant but used machine availability as a parameter to vary the level of complexity. The autonomy was varied by the number of work systems controlled by autonomous control methods. Philipp (2014) and Windt et al. (2010c) showed that the logistics performance improves with an increasing level of autonomy and decreases again with the highest autonomy level. The decreasing performance may lie in a chaotic behavior of the logistic objects, which is not in the sense of the overall goal when the degree of autonomy is too high. Scholz-Reiter et al. (2009c) also showed that for each complexity level, an increase of autonomy improved the logistic performance. This effect is even enhanced with rising complexity levels. Contrary to Philipp (2014) and Windt et al. (2010c), a decrease for the highest level of autonomy could not be observed. The different behavior of autonomous control in combination with a high autonomy level can probably be explained by the authors' different approaches for the definition of autonomy and complexity.

For transportation scheduling, various complexity levels for the vehicle-centric approach were compared (Wenning et al. 2006). It was shown that selecting routes solely based on packages locally available at a vehicle's starting point led to low utilization of vehicle capacities. The inclusion of packages, which can be loaded en route, significantly improved vehicle utilization and resulted in fewer vehicles needed for transport. The possibility of re-planning at each node, i.e., discarding the previous route and re-optimizing it on the basis of current data, was also investigated. The results showed that frequent rescheduling does actually impair the results compared to one-off route planning with consideration of the general cargo at other nodes. This can be explained by the fact that optimization was carried out for a complete route, of which only the first section was actually traveled before the rescheduling begins. This loss of performance could be partially compensated by a higher weighting of local general packages. Nevertheless, frequent rescheduling of routes can be disadvantageous.

For the $\mathrm{DLRP_t}$, it could be shown that measures to improve the scalability of information processing can significantly reduce communication volumes in large networks without impairing the logistical achievement of objectives. For that reason, the $\mathrm{DLRP_t}$ can be scaled to larger networks within the given limits of the communication infrastructure (Wenning et al. 2009).

So far, the results show that the autonomous control methods are suitable for production, transportation, and network scheduling and able to cope with complex environments and dynamic influences. Also, the results show that the methods' performance depends on the complexity level of the scenario. Furthermore, it was observed that the methods' performance depends on its characteristics (e.g., future- or past-oriented autonomous control methods).

5.2 Autonomous Control and Conventional Planning and Control

Autonomous control methods have been developed as an alternative to conventional planning and control methods. Thus, a profound evaluation of autonomous control requires a comparison between both autonomous control and conventional planning and control. This issue is addressed in different simulation studies. These studies mostly considered several scenarios with varying levels of dynamic influences. Scholz-Reiter et al. (2008a) compared conventional planning and control with the BEE algorithm in a shop-floor scenario and analyzed the impact of workstation breakdowns on the methods' performance. The results showed that the methods' performance was comparable without disturbances. In contrast, in the case of occurring workstation breakdowns, the BEE algorithm showed slightly better results. Scholz-Reiter and Jagalski (2008) used a 3 × 3 shop-floor scenario to compare conventional planning and control with the PHE method. The autonomous control method reached a better performance and seemed to be more robust against dynamic influences. Scholz-Reiter et al. (2009d) compared autonomous control with conventional planning and control in a production network scenario. An autonomously controlled network seemed to react more robustly to both changes in the order arrival rate and variations of the transport interval between the plants. Whereas the previously mentioned studies used rather simple methods for conventional planning and control, Scholz-Reiter et al. (2010a) applied more sophisticated methods, including several constructive heuristics and a genetic algorithm introduced by Jungwattanakit et al. (2009). The study considered different dynamic levels using a Poisson process with a varying number of released jobs per time unit. The results showed that the conventional planning methods performed best in static environments, whereas the performance of autonomous control improved with increasing dynamics. In highly dynamic environments, autonomous control outperformed conventional planning.

For transport scheduling, the DLRP_t was compared with benchmark solutions for the dynamical Vehicle Routing Problem (dynVRP) and the dynamic Pickup and Delivery Problem with Time Windows (dynPDPTW), like Tabu-Search and Mixed Integer Programming (MIP). The result showed that, for complex dynamical Vehicle Routing Problems with large problem sizes, the DLRP_t outperforms classical solutions (Rekersbrink et al. 2009; Scholz-Reiter et al. 2009a). Regarding the dynPDPTW problem, the performance achieved by the DLRP_t was comparable, and in some cases superior, to the benchmark solutions. In addition, compared to classical optimization approaches, the DLRP_t offers extended degrees of freedom and variability, e.g., object individual decision functions (Rekersbrink and Wenning 2011).

Schukraft et al. (2015) applied different coupling strategies for the combined application of conventional planning and autonomous control methods in a shop-floor scenario similar to Scholz-Reiter et al. (2010a). Depending on the extent of autonomous control for decision-making, these strategies can be rated from a low to a high level of autonomy. The performance measurement considered both logistic objectives achievement and adherence to predefined planning parameters. For production planning, several constructive heuristics have been used, introduced by Jungwattanakit et al. (2009). For autonomous control, the QLE and the PHE algorithms in combination with different priority rules have been applied. The simulation study considered different levels of dynamic influences to varying the complexity of the shop-floor scenario. The results show that there is no dominant coupling strategy for all scenarios and performance measurements. Naturally, high adherence to planning parameters supports a high planning accuracy. Especially strategies with autonomous order sequencing reach high logistic objectives achievement. In high dynamic environments, coupling strategies with a higher level of autonomy mostly outperform those with less autonomy. In contrast, in static environments, strategies with high adherence to preplanned parameters reach better results. These results could also be confirmed in a similar simulation study based on real data (Schukraft et al. 2016b).

Generally, the comparisons of autonomous control and conventional planning and control methods confirm the initial hypothesis that autonomous control methods are especially promising in high dynamic environments. In contrast, conventional planning and control methods reach comparable or even better results in static environments. Although some of the results are based on very simple conventional planning and control methods, the results could also be confirmed with more sophisticated conventional planning methods. Furthermore, the coupling of conventional planning and autonomous control showed that the performance also depends on the level of autonomy.

6 Application of Autonomous Control

So far, the presented results were mostly based on the theoretical reference scenarios described in Sect. 3.1. These scenarios are well suitable for the development and evaluation of autonomous control methods. However, they do not fully consider influencing factors that occur in practical application. Thus, to evaluate and prove the applicability of autonomous control in practice, the methods have also been applied in several real-world use cases described in the following.

6.1 Finished Vehicle Logistics

Finished vehicle distribution logistics covers all distribution processes of ready-made cars (and other vehicles) from the original equipment manufacturers (OEM) to the end customers, involving the OEM, various logistic service providers (LSP), and car dealers (Scholz-Reiter et al. 2012). It is an important branch within logistics, as many millions of cars have to be transported each year, a large part of them across countries and continents. The so-called vehicle compounds situated at focal points within the finished vehicle distribution networks, e.g., seaports, serve as transshipment and service points. At large compounds, several million vehicles per year can change the means of transport (trucks, railways, and ships). In addition, vehicles may be stored there intermediately for weeks or months, and technical operations may be performed, including washing, removing protective coats, installing luxury equipment, or technical modifications. At any given time, there can be up to 100,000 vehicles on a large compound (Klug 2018). For each of these vehicles, arrival (unloading), parking (storage), technical service operations, and finally departure (loading) have to be planned and executed. Thus, vehicle compound operators face difficult, complex, and highly dynamic planning and control tasks, rendering them a worthwhile application for autonomous control.

Consequently, a mechanism for parking space allocation was developed, which considers both the parking vehicles and the parking spaces as fully autonomous objects. Both object types can negotiate with each other and render their own decisions based on their local sets of objectives. The goal of the parking spaces is the highest possible occupation, and thus, unoccupied parking spaces will offer parking services for requesting vehicles. The objectives of the vehicles are to minimize their overall travel times on the compound. Therefore, a vehicle will choose the parking space that offers the shortest possible travel time from the unloading point to the parking space and from there to the subsequent destination (either a service station or the allocation area to the subsequent transport units). A negotiation mechanism regulates the interaction between the vehicles and the parking areas as follows:

1. Request of the vehicle for parking space, including its starting location and its destination.
2. Computation of distances and resulting travel times to an unoccupied parking area, resulting in an offer by that parking area that includes the travel time.
3. Comparison of all offers from different unoccupied parking areas by the requesting vehicle and selection of the parking area with the minimum travel time.

This negotiation mechanism was implemented using a multi-agent system. A simulation study was applied to compare the mechanism's performance to a central parking assignment based on fixed priority rules. The simulation study resulted in a reduced average vehicle travel time, equating to a time saving of 112 person workdays over a year for the studied compound (Böse et al. 2009; Böse and Windt 2007a). The improvements were caused by the capability to select the best available out of several alternative parking area offers.

Later simulation studies extended the scope of the autonomously controlled scenario by adding optional passage of the vehicles through a technical service center. In addition, different methods from autonomous control and related fields (e.g., $\mathrm{DLRP_t}$, pheromone Holonic, minimum buffer, QLE) were compared to conventional methods, e.g., predefined lists, corresponding to how cars were actually distributed on the studied compound. The simulations compared average travel times to parking and average parking area utilization for the storage processes. For the manufacturing processes, average total operation times of the vehicles and average due date reliability have been used. The results showed that the autonomous control methods could outperform the conventional methods under certain conditions, but no method dominates all others uniformly. Consequently, different areas of application exist for specific methods. In addition, preferences for different logistic objectives influence the selection of an appropriate autonomous control method. For instance, methods aiming at minimization of traveling times may cause low utilization of the parking areas. In contrast, other methods may cause opposite effects (Becker and Windt 2011; Windt et al. 2010b).

6.2 Apparel Logistics

As another application scenario, apparel distribution logistics was studied. Apparel logistics face conflicting demands of short delivery times, high product, and variant diversity, as well as uncertain delivery times due to the spatial separation of manufacturing in developing countries and customers in industrialized countries. For short-lived fashion products, seasonal order and delivery cycles are adopted, where many small-sized orders by individual retailers are aggregated into a large production lot that is delivered several times a year. Due to the long delivery times, which can extend to several months, some customer orders have to be anticipated using forecasts, while some orders that were initially placed are retracted or become obsolete. Consequently, some re-allocation of ready-made goods between different customer orders is necessary during distribution (Ait-Alla et al. 2014). Standardized routing causes bottlenecks in the distribution centers, including short-term peaks of stocks and workload.

In this regard, a mechanism for dynamic allocation of bundles of clothing articles, which are located at different locations within the distribution network, to customer orders was developed. Within each bundle, the individual garments could be identified using radiofrequency identification (RFID), and the composition of the bundle be compared to that of a customer order. Thus, restrictions regarding the ordered article variants (like different sizes and colors), quantities, and delivery dates can be taken into account (Scholz-Reiter et al. 2011a). The individual identification of goods and the allocation of goods packages to orders were implemented in a prototypical hardware and software application. The developed procedure was assessed by process simulation with regard to the achievable process potentials and by an economic efficiency analysis in terms of its economic effects (Scholz-Reiter et al. 2009b). Applying autonomous control to apparel logistics, however, proved to be difficult. In particular, a large number of products, each of limited value, limited technologies that could be attached (Scholz-Reiter et al. 2011a).

6.3 Event Logistics

Another application scenario was event management. Event management is concerned with the temporary renting of diverse equipment for use during organized events (e.g., public music concerts), its timely provision at the outside event venues, and setup in due time before the start of the event, as well as its dismantling and removal after the end of the event. Logistic processes associated with event management include the commissioning and transportation of highly diverse event equipment between venues, with strict deadline restrictions. They are characterized by a high level of dynamics and complexity (Harjes and Scholz-Reiter 2013b). In particular, order scheduling, which combines the allocation of resources to orders and the planning of cargo composition and transport routes for transports to, from, and between venues, seems an appropriate application area for autonomous control methods. Unexpected occurrences require frequent rescheduling of resources. The associated planning and control processes, particularly scheduling, require a correspondingly high degree of flexibility (Harjes and Scholz-Reiter 2013a). In response, an autonomous controlled scheduling system was developed. The scheduling is executed through mutual negotiations of autonomous objects represented by software agents, using the PlaSMA multi-agent simulation platform. A PlaSMA installation was set up that can transfer the current order situation of an event management company into a scenario, process simulations, and then transform the simulation results into planning decisions, generating event-specific picking and loading lists and a transport route and personnel schedules.

Within the PlaSMA installation, corresponding agents have been created to represent articles, vehicles, and employees. To start a planning cycle, the so-called list agents are created for advertised events, which coordinate the disposition of the resources for their respective events. For this purpose, they are equipped with event data like location, duration, and requested equipment. To start a negotiation, the list agent sends a request to all individual items of the types of equipment requested for the respective event. The requested item agents check the availability of their items for the event duration (they might already be scheduled for another parallel event or maintenance). Agents of available items send a request to the vehicle agents to clarify their movement to the event location. The vehicle agents, in turn, check whether, and at what cost, transport of the respective item to the destination is possible and then send corresponding offers or rejections to the items. For the generation of route plans, the DLRP_t protocol was implemented into the PlaSMA environment as a behavioral scheme of the transport agents (Harjes and Scholz-Reiter 2013a). This protocol allows the transport agents to find routes in a closed transportation network that is dynamically variable due to the only temporary existence of the event locations.

In a second turn, corresponding offers or rejections are transmitted in reverse order by the item agents to the initiating list agent. The list agent then selects the best of all offers it has received from items. The negotiation is complete when a list agent has been able to allocate the requested resources to its event. When the negotiation cycles have been completed for all list agents, all events are planned, and the simulation ends. If the scheduling cannot be completed with the available resources, the project manager must allow the use of additional rental vehicles or items, which are taken into account in a new simulation run. This procedure can be continued until a solution to the scheduling problem has been found. The final allocation of resources to projects can take several iterations (Harjes and Scholz-Reiter 2013a).

To enable the consideration of items already located at event locations, a hardware module for the acquisition and processing of the necessary material flow data at the event locations was developed, using RFID technology to identify equipment articles. The module can recognize the direction of loading and transmit the data obtained to the disposition system in a processed form.

As a result, the system can reduce the workload of experts involved in central planning and increase the efficiency and robustness of the planning under dynamic conditions. The simulations showed that for the studied event management company, the developed autonomous control system could reduce the number and aggregate distance of trips by between 20,000 and 52,000 km per year (Harjes and Scholz-Reiter 2014a, b).

6.4 Material Supply for Production

The internal logistics within production systems often uses the Kanban method (Ōno 2013). However, a lack of coordination between order release and order execution may cause problems during peak periods with unusually high demand. An intelligent network of different logistics units can improve the coordination between release and execution of the order. Correspondingly, a concept for a cyber-physical logistics system was developed by researchers from BIBA, which uses adapted autonomous control methods to realize an efficient, demand-oriented material supply in production (dynamic milk run). The concept was implemented in cooperation with a manufacturer of gear wheels, whose intra-logistics is based on a container-Kanban-procedure in combination with a milk run. In front of every machine, there is a delivery space where a one-floor roller (transport unit for several load carriers) for precisely one production order can be put down or picked up again. An employee picks up finished orders and delivers supplies to the machines with a tugger train. The machines are arranged so that the employee can reach all machines by driving an “eight” shaped course. At the intersection of both loops, the employee can optionally turn to the area for incoming and outgoing goods. Originally, the employee picked off finished orders, distributed them, and noted empty delivery areas in a fixed hourly turn. In the following cycle, he equipped these free delivery spaces with orders from the buffer stock. Since the employee did not have any information about the current collection and delivery orders, he drove all loops, although there was no need for transport in some loops. The fixed hourly cycle times meant that machines spent some idle time before new orders arrived, and the tugger train was not always fully utilized (Thoben et al. 2014).

The development of a cyber-physical logistics system focused on improved synchronization of the milk run with the actual demand through the networking of individual logistics units. Demand-driven supply of materials needs ongoing information about the current occupancy of delivery and pickup spaces.

For communication of stock levels, load carriers were equipped with sensors and a communication module, turning them into cyber-physical load carriers. The sensors could locate the load carriers and monitor environmental conditions (e.g., temperature, acceleration) affecting the components. The communication module could transmit delivery and pickup space-related data to a mobile data processing unit, which was provided to the tugger train driver. It ran software using the transmitted data to determine (remaining) machine processing times and completion dates of the current production orders in process and calculate the optimal departure time for a new delivery and pickup cycle.

The described cyber-physical logistics system matches autonomous control concepts with Industry 4.0 related technologies and thus emphasizes the close connection between both concepts as stated in Sect. 2. The potential of such a cyber-physical logistics system was assessed by a simulation study. The simulation results illustrated that the use of cyber-physical systems for a demand-driven material supply leads to a reduction of driven loops and an apparent reduction of driven cycles, compared to the previous situation. As a demand-driven material supply may lead to an increase of transport orders in a cycle, increasing the capacity of the tugger train may lead to a stronger reduction of the driven loops (Lappe et al. 2014; Thoben et al. 2014).

7 Conclusion and Outlook

7.1 Conclusion

This chapter summarized the research activities of BIBA in the development of autonomously controlled production systems and logistics networks that started in 2004 with the Collaborative Research Center 637 "Autonomous Cooperating Logistic Processes: A paradigm Shift and its Limitations" (cf. Sect. 1). In this context, this chapter presented autonomous control methods that have been developed for both autonomously controlled production systems and transportation networks. The first evaluation based on theoretical reference scenarios showed that autonomous control in large parts meets the expectations placed in it. This will briefly be explained, referring to the hypotheses regarding the performance of autonomous control (see Sect. 2) and the results of the theoretical evaluation (see Sect. 5).

The autonomous control methods have been applied in shop-floor, production network, and transportation scenarios with different types and intensities of complexity and dynamics. Often considered influences have been, e.g., fluctuating order arrivals, workstation breakdowns, or different network sizes. The results showed that autonomous control is able to deal with such influences and still maintain high performance (cf. hypothesis 1). Also, it could be shown that the methods' performance depends on the levels of complexity and autonomy. Naturally, a rising complexity leads to a decreasing performance (cf. hypothesis 3). However, in most cases, autonomous control is able to outperform conventional planning and control methods in environments of high complexity and dynamics. Contrarily, conventional planning and control methods reach comparable or even better results in static environments (cf. hypothesis 2). Although some of these results are based on a comparison of autonomous control with very simple conventional planning and control methods, the results could also be confirmed with more sophisticated conventional planning and control methods. Furthermore, approaches for the coupling of conventional planning methods with autonomous control showed that a rising level of autonomy increases the efficiency of production planning and control. Furthermore, a coupled application of conventional and autonomous control is promising to combine high efficiency with higher planning accuracy than that offered by autonomous control without consideration of conventional planning (cf. hypothesis 4).

In addition to the simulation-based evaluation, autonomous control was also applied in several application scenarios based on real-world use cases (see Sect. 6). These applications generally confirm the theoretical results and show that autonomous control has a high potential to improve the efficiency of planning and control in practice. For some application areas, however, autonomous control methods may not offer the expected performance improvements, or, e.g., in apparel logistics, may be too expensive to implement.

To summarize, it can be said that the hypotheses stated in Sect. 2 could be proven. However, production systems and transportation networks, in reality, are influenced by several factors of complexity and dynamics that could not be completely covered in the applied scenarios. Further research needs to consider further influences of complexity on the methods' performance and analyze the interdependencies between different influencing factors. In addition, the theoretical scenarios also neglect influences that are crucial for practical application and should be addressed in further research. This includes, e.g., the consideration of the influence of internal transport for material supply and the transport of products between different shops and workstations.

7.2 Outlook

The ongoing development of Industry 4.0 goes along with the anticipated improvements in ICT necessary for the implementation of autonomously controlled systems. Autonomous control and Industry 4.0 are supplemental concepts (cf. Sect. 2): Industry 4.0 provides the technological basis necessary for the decision-making of autonomous logistic objects, whereas autonomous control contributes the decision logic in the form of methods and algorithms to increase efficiency in logistics systems.

Among the real-life application examples of autonomous control (cf. Sect. 6), the development of a cyber-physical logistics system for material supply is particularly suited to illustrate the synergy effects between autonomous control and Industry 4.0. In addition, some of the developed autonomous control methods use a representational form of autonomous control, where physical logistic objects are represented by a digital object within a software system. Different types of logistic objects are represented by different software agents, using a MAS infrastructure for interaction. Examples are the negotiation between packages and transport units within the transportation scenario (cf. Sect. 4.2.2) and the negotiation between vehicles and parking spaces in the vehicle compound application (cf. Sect. 6.1). This virtual representation of physical objects transitioned into the concept of the digital twin within Industry 4.0 (Negri et al. 2017).

Until now, the research in autonomously controlled systems in the context of Industry 4.0 at BIBA was steadily continued. For example, current research work in the field of vehicle compound operations for finished vehicle logistics focuses on a combined optimal control system for order assignment and shuttle routing. Shuttles are used on vehicle compounds to transport handling employees to their subsequent order. A control algorithm assigns driving orders to handling employees and transport orders to shuttles and coordinates the routing of the shuttle buses (Hoff-Hoffmeyer-Zlotnik et al. 2020). The technical implementation uses both satellite-based and indoor positioning systems to locate vehicles, as well as WLAN connectivity. Another application concerns the development of cyber-physical transport bins that can monitor components being delivered within automotive supply chains. The transport bins are equipped with mobile wireless sensors and, via mobile communication technologies, connected to digital services hosted in cloud computing systems (Teucke et al. 2018). While the cyber-physical transport bins cannot render decisions on their own, they do possess certain forms of intelligence or information processing capability.

The continuing development of Industry 4.0 related technologies like cyber-physical systems and the Internet of Things and their introduction into the industrial application will make logistic objects progressively more intelligent and connected and bring new opportunities to realize autonomous control in manufacturing and logistic processes.

Acknowledgments The research presented in this chapter was funded by German Research Foundation (DFG) as part of the Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes—A Paradigm Shift and its Limitations” and the research project “Methods for the interlinking of central planning and autonomous control in production,” by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the project “SaSch—Digital Services for Shaping Agile Supply Chains,” by the German Federal Ministry of Education and Research (BMBF) within the project “Cyber Physical Production Systems—Enhancement of Productivity and Flexibility by Networking Intelligent Systems in the Factory,” and by the German Federal Ministry of Transport and Digital Infrastructure (BMVI) within the project “ISABELLA—Automobile logistics in sea- and inland ports: interactive and simulation-based operation planning, dynamic and context-based control of device- and load movements.”

References

Ait-Alla, A., Teucke, M., Lütjen, M., Beheshti-Kashi, S., Karimi, H.R.: Robust production planning in fashion apparel industry under demand uncertainty via conditional value at risk. Math. Probl. Eng. **2014**, 1–10 (2014). https://doi.org/10.1155/2014/901861

Alkan, B., Vera, D.A., Ahmad, M., Ahmad, B., Harrison, R.: Complexity in manufacturing systems and its measures: a literature review. Eur. J. Ind. Eng. **12**, 116–150 (2018). https://doi.org/10.1504/EJIE.2018.089883

Armbruster, D., de Beer, C., Freitag, M., Jagalski, T., Ringhofer, C.: Autonomous control of production networks using a pheromone approach. Phys. A: Stat. Mech. Appl. **363**, 104–114 (2006). https://doi.org/10.1016/j.physa.2006.01.052

Becker, T., Windt, K.: A comparative view on existing autonomous control approaches: observations from a simulation study. In: Hülsmann, M., Scholz-Reiter, B., Windt, K. (eds.) Autonomous Cooperation and Control in Logistics, pp. 275–289. Springer, Berlin, Heidelberg (2011)

Becker, M., Wenning, B.-L., Görg, C.: Integrated simulation of communication networks and logistical networks. In: Modeling and Simulation Tools for Emerging Telecommunication Networks, pp. 279–287. Springer, Boston, MA (2006)

Böse, F.: Selbststeuerung in der Fahrzeuglogistik: Modellierung und Analyse selbststeuernder logistischer Prozesse in der Auftragsabwicklung von Automobilterminals. In: Schriftenreihe: Informationstechnische Systeme und Organisation von Produktion und Logistik, vol. 14. Gito, Berlin (2012)

Böse, F., Windt, K.: Autonomously controlled storage allocation on an automobile terminal. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics, pp. 351–363. Springer, Berlin, Heidelberg (2007a)

Böse, F., Windt, K.: Catalogue of criteria for autonomous control in logistics. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics, pp. 57–72. Springer, Berlin, Heidelberg (2007b)

Böse, F., Piotrowski, J., Scholz-Reiter, B.: Autonomously controlled storage management in vehicle logistics – applications of RFID and mobile computing systems. Int. J. RF Technol. **1**, 57–76 (2009). https://doi.org/10.1080/17545730802294452

Bouhai, Grayson, S.: Internet of Things – Evolutions and Innovations, vol. 4. John Wiley & Sons, Hoboken, NJ (2017)

Broy, M.: Cyber-Physical Systems: Innovation Durch Softwareintensive Eingebettete Systeme. Springer, Dordrecht (2010)

Chaouchi, H.: The Internet of Things: Connecting Objects to the Web. Wiley, Somerset (2013)

Duffie, N., Bendul, J., Knollmann, M.: An analytical approach to improving due-date and lead-time dynamics in production systems. J. Manuf. Syst. **45**, 273–285 (2017). https://doi.org/10.1016/j.jmsy.2017.10.001

Freitag, M., Herzog, O., Scholz-Reiter, B.: Selbststeuerung logistischer Prozesse – Ein Paradigmenwechsel und seine Grenzen. Industrie Management. **20**, 23–27 (2004)

Grundstein, S.: Fertigungsregelung flexibler Fließfertigungen und Werkstattfertigungen zur Einhaltung von Lieferterminen. GRIN Verlag, Munich (2017)

Grundstein, S., Schukraft, S., Scholz-Reiter, B., Freitag, M.: Coupling order release methods with autonomous control methods – an assessment of potentials by literature review and discrete event simulation. Int. J. Prod. Manag. Eng. **3**, 43–56 (2015a)

Grundstein, S., Schukraft, S., Scholz-Reiter, B., Freitag, M.: Evaluation system for autonomous control methods in coupled planning and control systems. Procedia CIRP. **33**, 121–126 (2015b). https://doi.org/10.1016/j.procir.2015.06.023

Grundstein, S., Schukraft, S., Freitag, M., Scholz-Reiter, B.: Planorientierte autonome Fertigungssteuerung: Simulationsbasierte Untersuchung der Planeinhaltung und der logistischen Zielerreichung. wt Werkstattstechnik online. **105**, 220–224 (2015c)

Grundstein, S., Freitag, M., Scholz-Reiter, B.: A new method for autonomous control of complex job shops – integrating order release, sequencing and capacity control to meet due dates. J. Manuf. Syst. **42**, 11–28 (2017). https://doi.org/10.1016/j.jmsy.2016.10.006

Harjes, F., Scholz-Reiter, B.: Agent-based disposition in event logistics. Res. Prod. Logist. **3**, 137–150 (2013a)

Harjes, F., Scholz-Reiter, B.: Selbststeuerndes Routing für Verleihartikel. Industrie Management. **29**, 44–48 (2013b)

Harjes, F., Scholz-Reiter, B.: Autonomous control in closed dynamic logistic systems. Procedia Technol. **15**, 313–322 (2014a). https://doi.org/10.1016/j.protcy.2014.09.085

Harjes, F., Scholz-Reiter, B.: Integration aspects of autonomous control in event logistics. Res. Logist. Prod. **4**, 5–20 (2014b)

Hermann, M., Pentek, T., Otto, B.: Design principles for Industrie 4.0 scenarios. In: 49th Hawaii International Conference on System Sciences, pp. 3928–3937 (2016)

Hoff-Hoffmeyer-Zlotnik, M., Sprodowski, T., Freitag, M.: Revisiting order assignment problems in a real-case vehicle compound scenario. In: 2020 6th IEEE Congress on Information Science and Technology, pp. 395–400. IEEE, Piscataway (2020)

Hon, K.K.B.: Performance and evaluation of manufacturing systems. CIRP Ann. **54**, 139–154 (2005). https://doi.org/10.1016/S0007-8506(07)60023-7

Jungwattanakit, J., Reodecha, M., Chaovalitwongse, P., Werner, F.: A comparison of scheduling algorithms for flexible flow shop problems with unrelated parallel machines, set-up times, and dual criteria. Comput. Oper. Res. **36**, 358–378 (2009). https://doi.org/10.1016/j.cor.2007.10.004

Kagermann, H, Wahlster, W., Helbig, J.: Umsetzungsempfehlung für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0. aca-tech – Deutsche Akademie der Technikwissenschaften e. V. (2013)

Klug, F.: Logistikmanagement in der Automobilindustrie. Springer, Berlin, Heidelberg (2018)

Lappe, D., Veigt, M., Franke, M., Kolberg, D., Schlick, J., Stephan, P., Guth, P., Zimmerling, R.: Vernetzte Steuerung einer schlanken Intralogistik: Simulationsbasierte Potentialanalyse einer bedarfsorientierten Materialversorgung in der Fertigung. wt Werkstattstechnik online. **3**, 112–117 (2014)

Negri, E., Fumagalli, L., Macchi, M.: A review of the roles of digital twin in CPS-based production systems. Procedia Manufacturing. **11**, 939–948 (2017). https://doi.org/10.1016/j.promfg.2017.07.198

Nyhuis, P., Wiendahl, H.-P.: Fundamentals of Production Logistics: Theory, Tools and Applications. Springer, Berlin (2009)

Ōno, T.: Das Toyota-Produktionssystem. Produktion. Campus-Verl, Frankfurt (2013)

Ouelhadj, D., Petrovic, S.: A survey of dynamic scheduling in manufacturing systems. J. Sched. **12**, 417–431 (2009). https://doi.org/10.1007/s10951-008-0090-8

Perkins, C.E.: Ad Hoc Networking: An Introduction. Addison-Wesley, Boston (2001)

Philipp, T.: Messung und Bewertung selbststeuernder Produktionssysteme. Doctoral dissertation, University of Bremen (2014)

Philipp, T., Böse, F., Windt, K.: Autonomously controlled processes-characterisation of complex production systems. In: Proceedings of 3rd CIRP Conference in Digital Enterprise Technology, Setubal, Portugal (2006)

Philipp, T., de Beer, C., Windt, K., Scholz-Reiter, B.: Evaluation of autonomous logistic processes – analysis of the influence of structural complexity. In: Understanding Autonomous Cooperation and Control in Logistics, pp. 303–324. Springer, Berlin, Heidelberg (2007)

Rekersbrink, H., Wenning, B.-L.: Selbststeuerung in der Transportlogistik – Evaluation einer Selbststeuerungsmethode anhand dynamischer Pickup and Delivery Probleme. Industrie Management. **27**, 30–34 (2011)

Rekersbrink, H., Ludwig, B., Scholz-Reiter, B.: Entscheidungen selbststeuernder logistischer Objekte. Industrie Management. **23**, 25–30 (2007)

Rekersbrink, H., Makuschewitz, T., Scholz-Reiter, B.: A distributed routing concept for vehicle routing problems. Logist. Res. **1**, 45–52 (2009). https://doi.org/10.1007/s12159-008-0003-4

Rekersbrink, H., Scholz-Reiter, B., Zabel, C.: An autonomous control concept for production logistics. In: Advanced Manufacturing and Sustainable Logistics, pp. 245–256. Springer, Berlin, Heidelberg (2010)

Scholz-Reiter, B., Jagalski, T.: Dynamics of autonomous control in production logistics. In: 26th International Conference of the System Dynamics Society, p. 15. System Dynamics Society, Albany (2008)

Scholz-Reiter, B., Freitag, M., de Beer, C., Jagalski, T.: Modelling and analysis of autonomous shop floor control. In: Proceedings of 38th CIRP International Seminar on Manufacturing Systems. Universidade Federal de Santa Catarina, Florianopolis (2005a)

Scholz-Reiter, B., Freitag, M., de Beer, C., Jagalski, T.: Modelling dynamics of autonomous logistic processes: discrete-event versus continuous approaches. CIRP Ann. **54**, 413–416 (2005b). https://doi.org/10.1016/s0007-8506(07)60134-6

Scholz-Reiter, B., Rekersbrink, H., Freitag, M.: Kooperierende Routingprotokolle zur Selbststeuerung von Transportprozessen. Industrie Management. **22**, 7–10 (2006a)

Scholz-Reiter, B., Freitag, M., de Beer, C., Jagalski, T.: Modelling and simulation of a pheromone based shop floor control. In: Proceedings of the 3rd International CIRP Sponsored Conference on Digital Enterprise Technology, Setubal, pp. 1–7 (2006b)

Scholz-Reiter, B., Freitag, M., de Beer, C., Jagalski, T.: The influence of production networks' complexity on the performance of autonomous control methods. In: Proceedings of the 5th CIRP International Seminar on Computation in Manufacturing Engineering, pp. 317–320. Naples, Italy (2006c)

Scholz-Reiter, B., Jagalski, T., de Beer, C., Freitag, M.: Autonomous shop floor control considering set-up times. In: Proceedings of the 40th CIRP International Seminar on Manufacturing Systems. University of Liverpool, Liverpool (2007a)

Scholz-Reiter, B., Kolditz, J., Hildebrandt, T.: UML as a basis to model autonomous production systems. In: Digital Enterprise Technology, pp. 553–560. Springer, Berlin (2007b)

Scholz-Reiter, B., Jagalski, T., Bendul, J.C.: Autonomous control of a shop floor based on Bee's foraging behaviour. In: Kreowski, H.-J., Scholz-Reiter, B., Haasis, H.-D. (eds.) Dynamics in Logistics, pp. 415–423. Springer, Berlin, Heidelberg (2008a)

Scholz-Reiter, B., de Beer, C., Freitag, M., Jagalski, T.: Bio-inspired and pheromone-based shop-floor control. Int. J. Comput. Integr. Manuf. **21**, 201–205 (2008b). https://doi.org/10.1080/09511920701607840

Scholz-Reiter, B., Rekersbrink, H., Wenning, B.-L., Makuschewitz, T.: A survey of autonomous control algorithms by means of adapted vehicle routing problems. In: Proceedings of the 9th Biennial ASME Conference on Engineering Systems Design and Analysis, pp. 411–416. American Society of Mechanical Engineers Digital Collection, New York (2009a)

Scholz-Reiter, B., Özsahin, M.E., Teucke, M., Mehrsai, A.: Autonomous order allocation in large distance apparel industry supply chains based on use of RFID technology. In: Proceedings of the 14th International Symposium On Logistics, pp. 329–336. Global Supply Chains and Inter-Firm Networks, Centre for Concurrent Enterprise Nottingham University Business School, Nottingham (2009b)

Scholz-Reiter, B., Görges, M., Philipp, T.: Autonomously controlled production systems - influence of autonomous control level on logistic performance. CIRP Ann. **58**, 395–398 (2009c). https://doi.org/10.1016/j.cirp.2009.03.011

Scholz-Reiter, B., Görges, M., Jagalski, T., Mehrsai, A.: Modelling and analysis of autonomously controlled production networks. IFAC Proc. **42**, 846–851 (2009d). https://doi.org/10.3182/20090603-3-RU-2001.0081

Scholz-Reiter, B., Rekersbrink, H., Görges, M.: Dynamic flexible flow shop problems – scheduling heuristics vs. autonomous control. CIRP Ann. **59**, 465–468 (2010a). https://doi.org/10.1016/j.cirp.2010.03.030

Scholz-Reiter, B., Zabel, C., Rekersbrink, H.: Influence of competition-based knowledge restrictions on autonomous routing in transport logistics. In: Proceedings of the 2nd International Conference on Logistics and Transport, Queenstown, Neuseeland (2010b)

Scholz-Reiter, B., Görges, M., Jagalski, T., Naujok, L.: Modelling and analysis of an autonomous control method based on bacterial chemotaxis. In: 43rdCIRP International Conference on Manufacturing Systems 2010, pp. 699–706. Neuer Wissenschaftlicher Verlag, Wien (2010c)

Scholz-Reiter, B., Ruthenbeck, C., Teucke, M., Hoppert, J.: Limitations of autonomous control in practical applications: report on lessons learned from vehicle and apparel logistics. In: Autonomous Cooperation and Control in Logistics, pp. 291–311. Springer, Berlin, Heidelberg (2011a)

Scholz-Reiter, B., Görges, M., Jagalski, T..: Modeling and analyzing integrated autonomously controlled production and transport processes. In: Proceedings of the 21st International Conference On Production Research, vol 21 (2011b)

Scholz-Reiter, B., Görges, M., Rekersbrink, H.: Potentials and limitations of autonomously controlled production systems. In: Autonomous Cooperation and Control in Logistics, pp. 131–147. Springer, Berlin, Heidelberg (2011c)

Scholz-Reiter, B., Dashkovskiy, S., Görges, M., Naujok, L.: Stability analysis of autonomously controlled production networks. Int. J. Prod. Res. **49**, 4857–4877 (2011d). https://doi.org/10.1080/00207543.2010.505215

Scholz-Reiter, B., Werthmann, D., Lappe, D., Otterstedt, N.: Ortungsgestützte Steuerung von Qualitätssicherungsprozessen. Industrie Management. **28**, 22–26 (2012)

Schukraft, S., Grundstein, S., Freitag, M., Scholz-Reiter, B.: Strategies for the coupling of autonomous control and central planning: Evaluation of strategies using logistic objectives achievement and planning adherence. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 121–126. IEEE, Piscataway (2015)

Schukraft, S., Grundstein, S., Scholz-Reiter, B., Freitag, M.: Evaluation approach for the identification of promising methods to couple central planning and autonomous control. Int. J. Comput. Integr. Manuf. **29**, 438–461 (2016a). https://doi.org/10.1080/0951192X.2015.1066032

Schukraft, S., Grundstein, S., Freitag, M., Scholz-Reiter, B.: Evaluation of strategies for the coupling of central planning and autonomous control in dynamic job shop environments. Adv. Mater. Res. **1140**, 457–464 (2016b). https://doi.org/10.4028/www.scientific.net/AMR.1140.457

Singh, G., Wenning, B.-L., Görg, C.: Efficient communication in autonomous logistic processes by application of cluster-based routing methods. In: Proceedings of Summer Computer Simulation Conference (2008)

Sowade, S., Rippel, D., Scholz-Reiter, B.: Modeling concept for the infrastructure of autonomous logistic processes. CIRP J. Manuf. Sci. Technol. **5**, 254–266 (2012). https://doi.org/10.1016/j.cirpj.2012.09.011

Teucke, M., Broda, E., Börold, A., Freitag, M.: Using sensor-based quality data in automotive supply chains. Mach. Des. **6**, 53 (2018). https://doi.org/10.3390/machines6040053

Thoben, K.-D., Veigt, M., Lappe, D., Franke, M., Kück, M., Kolberg, D., Fahl, I., Zimmerling, R., Schlick, J., Stephan, P., Guth, P.: Towards networking logistics resources to enable a demand-driven material supply for lean production systems – basic concept and potential of a cyber-physical logistics system. In: Proceedings of the 7. BVL Scientific Symposium on Logistics, pp. 42–69. BVL International Bundesvereinigung Logistik (BVL) e.V, Bremen (2014)

University of Bremen.: Bremen Research Cluster for Dynamics in Logistics (LogDynamics). http://www.logdynamics.de/logdynamics.html?&L=1 (2021). Accessed 18 June 2021

Warden, T., Porzel, R., Gehrke, J.D., Herzog, O., Langer, H., Malaka, R.: Towards ontology-based multiagent simulations: the Plasma approach. In: Bargiela, A. (ed.) 24th European Conference on Modelling and Simulation ECMS 2010: June 1st–4th, 2010, Kuala Lumpur, Malaysia: Proceedings, pp. 50–56. European Council for Modelling and Simulation (ECMS), Dudweiler (2010)

Wenning, B.-L.: Context-Based Routing in Dynamic Networks. Vieweg+Teubner, Wiesbaden (2010)

Wenning, B.-L., Görg, C., Peters, K.: Ereignisdiskrete Modellierung von Selbststeuerung in Transportnetzen. Industrie Management. **21**, 53–56 (2005)

Wenning, B.-L., Pesch, D., Timm-Giel, A.: A distributed routing approach for vehicle routing in logistic networks. In: 64th IEEE Vehicular Technology Conference: Montreal, Quebec, Canada. IEEE, Piscataway, NJ (2006)

Wenning, B.-L., Rekersbrink, H., Becker, M., Timm-Giel, A., Görg, C., Scholz-Reiter, B.: Dynamic transport reference scenarios. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics, pp. 337–350. Springer, Berlin, Heidelberg (2007a)

Wenning, B.-L., Görg, C., Timm-Giel, A., Schönberger, J., Kopfer, H.: Investigations on object-centered routing in dynamic environments: algorithmic framework and initial numerical results. In: Proceedings of the 9th International Conference on Enterprise Information Systems, pp. 225–230. Madeira, Portugal (2007b)

Wenning, B.-L., Rekersbrink, H., Görg, C.: Scalability investigations on communication traffic in distributed routing of autonomous logistic objects. In: 9th International Conference on Intelligent Transport Systems Telecommunications: Lille, France, pp. 8–12. IEEE, Piscataway, NJ (2009)

Wenning, B.-L., Pesch, D., Timm-Giel, A., Görg, C.: Environmental monitoring aware routing: making environmental sensor networks more robust. Telecommun. Syst. **43**, 3–11 (2010). https://doi.org/10.1007/s11235-009-9191-8

Wenning, B.-L., Rekersbrink, H., Timm-Giel, A., Görg, C.: Weighted multiplicative decision function for distributed routing in transport logistics. In: Dynamics in Logistics, pp. 117–124. Springer, Berlin, Heidelberg (2011)

Windt, K., Hülsmann, M.: Changing paradigms in logistics — understanding the shift from conventional control to autonomous cooperation and control. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics, pp. 1–16. Springer, Berlin, Heidelberg (2007)

Windt, K., Philipp, T., Böse, F.: Complexity cube for the characterization of complex production systems. Int. J. Comput. Integr. Manuf. **21**, 195–200 (2008). https://doi.org/10.1080/09511920701607725

Windt, K., Becker, T., Jeken, O., Gelessus, A.: A classification pattern for autonomous control methods in logistics. Logist. Res. **2**, 109–120 (2010a). https://doi.org/10.1007/s12159-010-0030-9

Windt, K., Becker, T., Kolev, I.: A comparison of the logistics performance of autonomous control methods in production logistics. In: Proceedings of the 43rd CIRP International Conference on Manufacturing Systems. Sustainable Production and Logistics in Global Networks, pp. 576–583. NW Verlag, Vienna (2010b)

Windt, K., Philipp, T., Böse, F., Becker, T.: Application of a three-component evaluation system for autonomous control in logistics. Proc. Inst. Mech. Eng. B J. Eng. Manuf. **224**, 1267–1276 (2010c). https://doi.org/10.1243/09544054JEM1656

Explorable Uncertainty Meets Decision-Making in Logistics

Nicole Megow and Jens Schlöter

Abstract Decision-making under uncertainty is a major challenge in logistics. Mathematical optimization has a long tradition in providing powerful methods for solving logistics problems. While classical optimization models for uncertainty in the input data do not consider the option to actively query the precise value of uncertain input elements, this option is in practice often available at a certain cost. The recent line of research on optimization under explorable uncertainty develops methods with provable performance guarantees for such scenarios. In this chapter, we highlight some recent results from the mathematical optimization perspective and outline the potential power of such model and techniques for solving logistics problems.

1 Uncertainty in Logistics

Uncertainty is a major challenge in effective decision-making in logistics and supply chain management (Wilding 1998). Uncertainty may refer to any lack of information about the supply chain, its environment or conditions, and the unpredictable impact of decisions (van der Vorst and Beulens 2002). There are numerous sources of uncertainty and dynamics in the system parameters; for a classification and survey, we refer to Simangunsong et al. (2012), Sanchez-Rodrigues et al. (2010). As illustrating examples consider varying transportation times depending on traffic conditions, weather, or disruptions; variable demand; deteriorating conditions of delivered goods; changing production capacity or speed due to aging or replacement of production units, etc.

In addition, digitalization and data-driven applications give permanent access to large amounts of data that can be used for planning and decision-making purposes.

N. Megow (✉) · J. Schlöter
Faculty of Mathematics and Computer Science, University of Bremen, Bremen, Germany
e-mail: nicole.megow@uni-bremen.de; jschloet@uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_2

However, these data need to be analyzed and processed to provide valuable insights, and moreover, they increase the level of dynamics and uncertainty significantly.

Mathematical optimization has a tradition in providing powerful methods for effectively planning logistics processes, such as routing and delivery, replenishment, and scheduling.

Due to the ubiquity of uncertainty (in fact, in all kinds of real-world applications), a plethora of research articles is devoted to decision-making under data uncertainty and optimization methods that are capable of handling large amounts of dynamically changing data. From the mathematical optimization perspective, there are three major frameworks for modeling uncertainty in the input data: stochastic optimization, online optimization, and robust optimization. Below, we sketch their general characteristics without reviewing the numerous individual settings and the respective literature.

Stochastic optimization generally refers to settings in which there is some randomness in the data. Typically, parts of the input data are modeled as random variables that follow a given probability distribution. Such probabilistic information may stem from statistical information, such as, travel time distributions from traffic surveillance or customer statistics. The goal is to find a solution that is (approximately) optimal in expectation. Different models define how adaptively an algorithm may proceed in making decisions and in what manner the true realization of uncertain data is revealed. Typically, we distinguish single-stage and multi-stage problems, but there are more dynamic models; see, e.g., the textbook by Birge and Louveaux (2011).

Online optimization refers to settings in which parts of the input are completely unknown a priori. The data is revealed incrementally one by one or over time, and an algorithm must make irrevocable decisions given only partial information; see, e.g., the book by Borodin and El-Yaniv (1998). As an example, consider the (often) unpredictable arrival of maintenance requests and a service provider who must immediately allocate and coordinate resources for maintenance and possibly replacement. It is typically impossible to give an algorithm that finds an optimal solution for all possible online inputs. Therefore, we evaluate the performance of algorithms by *competitive analysis*, a worst-case analysis that compares (implicitly), for any possible input instance, the cost of an algorithm's solution with the optimal cost.

Robust optimization is a framework in which, again, there are no probabilistic assumptions made. Typically, a set of possible input scenarios is given either explicitly or implicitly, e.g., by giving uncertainty intervals in which the true value of some unknown data lies. In the classical setting, we ask for a single solution that performs well in any possible input scenario. Hence, algorithms are absolutely non-adaptive and make decisions without any chance of adjusting them later. More recently, models have been proposed that allow some recovery actions once the scenario has been realized; see, e.g., the book by Ben-Tal et al. (2009).

In all these traditional frameworks for data uncertainty, optimization methods have to accept the incompleteness of input data. They can rely only on algorithmic intelligence to cope with it since there is *no way to explicitly query for precise* *data* without committing to a partial solution. Clearly, more information or even knowing the exact data would allow for significantly better solutions. The possibility of querying is a reasonable modeling assumption as many real-world applications provide the possibility to obtain exact data at a certain exploration cost such as extra time, money, bandwidth, etc.

The framework of **explorable uncertainty** captures problem settings in which, for any uncertain input element, a *query* can be used to obtain the exact value of that element. Queries are costly, and hence the goal is to make as few queries as possible until sufficient information has been obtained to solve the given problem. The major challenge is to balance the resulting exploration–exploitation trade-off.

Outline of the Chapter

The goal of this chapter is to highlight some recent results on explorable uncertainty from the mathematical optimization perspective and to outline the potential power of such framework and techniques for solving logistics problems. In the following sections, we introduce three models within this framework.

1. An *online model*, in which an uncertain input element can attain any value within a given interval.
2. A *learning-based model*, in which an algorithm has access to a prediction, possibly machine-learned, on the exact value of an uncertain element but without any guarantee on its accuracy.
3. A *stochastic model*, in which there is given probabilistic information about the exact value.

We give intuitive insights, state mathematical results, and outline methods to tackle problems under explorable uncertainty. To illustrate these insights, we consider the example problems of selecting a minimum element within a set of uncertain values by querying a minimal number of elements (*minimum problem*) as well as a query-based variant of the *minimum spanning tree (MST) problem*, a most fundamental network design problem that asks for the minimum cost network (a cycle-free connected graph). These are important problems that appear as subproblems in many applications, also in logistics, and they have been studied extensively in the framework of explorable uncertainty, offering positive results in all three models.

2 Power of Exploring Uncertain Data in Logistics

We showcase different logistics scenarios in which methods building on explorable uncertainty promise a substantial improvement of the current state of the art.

Consider the task of designing an effective complex supply chain (or certain parts of it) for a company with various production facilities, sub-contractors, storage facilities, and customers. At several places in the decision-making process, the management has to decide for a best choice out of a number of choices where some parameters might be uncertain. Such choices could be possible sub-contractors, facility locations, transportation means, etc. Which is the "best" may depend on parameters such as cost, quality, reliability, or customer satisfaction and might not be known, but there is the possibility to query those parameters by, e.g., measurements, lab tests, customer interviews, etc. This is typically costly and shall be used only if the effect on the performance of supply chain processes is worth it. Powerful and well-thought exploration methods may give a substantial improvement over decision-making under uncertainty.

The next example falls into the area of transport logistics of *perishable goods*, that is, goods with a decay time in the same order of magnitude as the transport time, e.g., fruits, vegetables, meat, or cheese in sea and land transportation. A major challenge in the transportation of perishable goods is the uncertainty in the state of the goods and their environment, e.g., in a container. The ripeness degree and shelf life of fresh fruit are hard to predict as they may change drastically and very fast. During a typical transport, e.g., maritime container transport or transportation via truck or train, there is no feedback on the current state of the goods and no timely action to avoid food loss can be taken. With an *intelligent container* (a keyword coined by Lang and Jedermann (2016)), i.e., a container equipped with sensors and infrastructure for monitoring, communication, and possibly even further actions, such an early feedback and even active queries for more precise information are possible and could be integrated into an adaptive logistics planning framework.

More concretely, consider a truck routing problem for delivering fresh food to several possibly widespread destinations, e.g., raspberries from Turkey delivered by truck to several destinations in Germany. Clearly, frequent sensor measurement and resulting updates on the remaining shelf time cost energy and decrease the life time of a battery. Replacing and disposing batteries costs time, money, and it requires the expertise (of the truck driver) to do so. This may not be relevant for a single 3-day trip, but it might play a significant role as trucks drive tours repeatedly. Dynamic routing algorithms that incorporate data exploration may minimize the food loss while not exceeding a given budget of exploration cost measured, e.g., in energy.

Another serious problem in container transport is pest insects and dangerous gases, which are detected only when unloading at a port where ad hoc protective measures are costly and inefficient. Consider the scheduling and resource allocation for handling containers that require special treatment. The operations for opening dangerous containers in a secure area (chemicals and gas) or rerouting a container (pest insects) are time- and resource-expensive. Suppose sensor measurements can be inquired, and the relevant information can be made available on time for efficiently planning the necessary port operations. It might be more efficient to open several dangerous containers in a secure area at the same time than blocking resources any time a single such container arrives. This is only possible if the information is queried at the right point in time while avoiding unnecessary query cost, e.g., energy.

Besides the technical possibility to query uncertain data, it is a major challenge to algorithmically balance the cost for data exploration (e.g., extensive tests in a lab, energy consumption for queries to the intelligent container) with the benefit for the quality of a solution (e.g., cost for establishing a network, amount of food loss).

3 Optimization Under Explorable Uncertainty

In explorable uncertainty, we are given, instead of precise data points, only rough information in form of *uncertainty intervals*. Precise data points can be revealed by using *queries*. Since querying data points comes at a cost, the goal is to extract sufficient information to solve the problem while minimizing the total query cost. In the following, we formally introduce the model of explorable uncertainty and highlight important concepts for solving problems under explorable uncertainty.

3.1 The Model

We are given a ground set $\mathcal{I}$ of uncertainty intervals. Each interval $I_i \in \mathcal{I}$ is associated with a precise value $w_i \in I_i$ that is initially unknown. The precise value of an uncertainty interval I_i can be extracted by using a query. Intuitively, querying the interval $I_i = (L_i, U_i)$ replaces the open interval (L_i, U_i) with the singleton $[w_i]$. We call L_i and U_i the lower and upper limits of I_i. How to obtain the upper and lower limits of the uncertainty intervals is problem specific and depends on the application. As an example consider the distances between mobile agents. While the agents change their positions and the precise distance between two agents might not always be known, last known locations as well as maximum movement speeds can be used to compute an uncertainty interval that is guaranteed to contain the precise distance. In the following, we abstract from the process of obtaining the uncertainty intervals and assume they are part of the input. If $I_i = [w_i]$, we define $L_i = U_i = w_i$. A query to an interval I_i comes with a *query cost* of c_i. For the remainder of this chapter, we only consider uniform query costs, i.e., $c_i = 1$ for all $I_i \in \mathcal{I}$.

We can define various optimization problems based on the ground set of uncertainty intervals. For each problem, the goal is to extract sufficient information to solve the problem for a fixed but initially unknown realization of precise values, while minimizing the total query costs. In the case of uniform query costs, the total cost is just the number of queried intervals. A *query set* $Q \subseteq \mathcal{I}$ is *feasible* if querying Q extracts sufficient information to optimally solve the problem at hand. Thus, a query set Q is only feasible if querying Q allows us to compute a solution for the underlying optimization problem that is guaranteed to be optimal for all possible precise values of intervals in $\mathcal{I} \setminus Q$. We further discuss this assumption at the end of the chapter. We analyze the performance of algorithms in terms of their *competitive ratio*. Let $\mathcal{J}$ denote the set of all instances for a problem under explorable uncertainty, let $\text{ALG}(J)$ for $J \in \mathcal{J}$ denote the query cost needed by an algorithm ALG to solve instance J, and let $\text{OPT}(J)$ denote the optimal query cost for solving J. That is, for a fixed instance J with fixed precise values, $\text{OPT}(J)$ denotes the minimum query cost necessary to solve J. Then, the competitive ratio of ALG is defined as

$$\max_{J \in \mathcal{J}} \frac{\text{ALG}(J)}{\text{OPT}(J)}.$$

In the following, we introduce two example problems under explorable uncertainty.

3.1.1 Example: Minimum and Selection Problems

In the *minimum problem*, the goal is to determine, for a given set of uncertainty intervals $\mathcal{I}$, an interval $I_i \in \mathcal{I}$ with minimum precise value, i.e., $I_i = \arg\min_{I_i \in \mathcal{I}} w_i$. Note that this problem does not necessarily involve computing the actual precise value of that interval.

As an example recall the scenario given in Sect. 2, where a company has to select the “best” out of a pool of possible sub-contractors, facility locations, transportation means, etc. without having all the information to determine it. This scenario can be modeled as a minimum problem: the possible choices can be modeled by the index set $\{1, \ldots, n\}$. For each possible choice $i \in \{1, \ldots, n\}$, we have an initial estimation of its quality (based, e.g., on publicly available information, past experiences, and already known basic conditions) that can be modeled by the uncertainty interval I_i. A precise estimation for a possible choice can be obtained, e.g., by executing measurements, lab tests, or customer interviews. Then, the process of obtaining a precise estimation for a possible choice can be modeled by a query. As the described operations come typically at a high cost, the goal is to make the best possible choice while minimizing this extra cost. This corresponds to the minimum problem.

Since the precise values are initially unknown, it might not be possible to find the interval of minimum precise value without executing queries. For example, in Fig. 1, we are given a set of two uncertainty intervals with the task to determine the interval with minimum precise value. Since those intervals overlap, both of them could possibly be of minimum precise value. To solve the problem, an algorithm has to execute at least one query.

The example of Fig. 1 also shows that no algorithm is better than 2-competitive for the minimum problem, as Kahan (1991) observed already in his seminal paper. By definition, for an algorithm to be better than 2-competitive, the ratio between $\text{ALG}(J)$ and $\text{OPT}(J)$ has to be strictly smaller than 2 for *every* instance J. In the example, we consider two instances with the same intervals that differ only

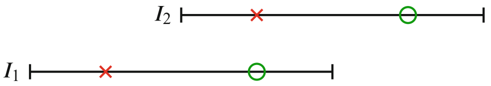


Fig. 1 Lower bound example for the minimum problem in a single set

in the precise values (crosses vs. circles). Since an algorithm has no knowledge of the precise values, both instances look the same to the algorithm, and thus, a deterministic algorithm will make the same first query for both instances. We argue that each possible first query will lead to a ratio of 2 for at least one of the instances, which implies that no deterministic algorithm is better than 2-competitive. In such a worst-case analysis, we may assume that different precise values are revealed for different algorithms. (In general, the precise values are independent of the query order.) If an algorithm queries I_1 first, then, in the worst case, the green circle is revealed as the precise value of I_1. After querying I_1, it is still unknown which interval has minimum value, which forces the algorithm to also query I_2. If the query to I_2 again reveals the green circle as the precise value of I_2, an optimal query set could determine that I_1 has minimum precise value by only querying I_2. Thus, the cost of the algorithm is twice the cost of the optimal query set. Vice versa, if an algorithm queries I_2 first, then, in the worst case, the red crosses are revealed as precise values, and the algorithm queries $\{I_1, I_2\}$, while the optimal query set queries only I_1. Hence, for any algorithm's choice on this instance (either query I_1 first or I_2), there is a realization of precise values on which the algorithm requires two queries, whereas an optimal query set with one query exists. This implies that no deterministic algorithm (an algorithm that makes the same decisions when given the same input) can be better than 2-competitive.

In a more general variant of the minimum problem, we are given a family $\mathcal{S}$ of (possibly non-disjoint) subsets of $\mathcal{I}$, and the goal is to determine the member of minimum precise value for each subset $S_j \in \mathcal{S}$. Consider the example given in Bampis et al. (2021) concerning a multi-national medical company. The company relies on certain products for its operation in each country, e.g., a chemical ingredient or a medicine. However, due to different approval mechanisms, the concrete products that are accessible differ for each country. The task is to find the best approved product for each country. The product quality can be determined by extensive tests in a lab (queries) and, since the quality is independent of the country, each product has to be tested at most once. The set of products available in one country corresponds to a set in $\mathcal{S}$, and the problem of identifying the best product in each country is the minimum problem in multiple sets.

In a similar way, we can model other selection problems, e.g., finding the kth smallest element and sorting.

3.1.2 Example: Minimum Spanning Tree Problem

In the *minimum spanning tree (MST) problem*, we are given a weighted, undirected, connected graph $G = (V, E)$, with nodes V and edges E, where each edge $e \in E$ has associated a weight $w_e \geq 0$. The task is to find a spanning tree of minimum total weight. A spanning tree is a connected acyclic graph whose edges span all the vertices. See Fig. 2 for an example graph. The MST problem has various applications, e.g., in the design of distribution networks: nodes can be used to model storage facilities, manufacturers, and transportation systems, while the edges and

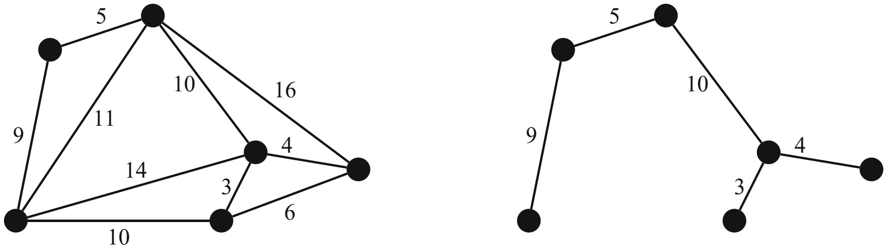


Fig. 2 Graph *without* uncertainty in the edge weights (left) and the corresponding minimum spanning tree (right)

their weights can model the cost of establishing a direct connection between two such points of interest, where a direct connection could, for example, be a road, a pipeline, or an Ethernet connection. To establish connections between all points of interest in a cost-minimal way, we have to compute a minimum spanning tree.

In the MST problem with uncertainty, the precise edge weights w_e are unknown. Each edge $e \in E$ is associated with an uncertainty interval $I_e \in \mathcal{I}$, and w_e is guaranteed to be in the given interval I_e. The task is to find an MST in the uncertainty graph G for an a priori unknown realization of edge weights. Note that this problem does not necessarily involve computing the actual MST weight. In the application given above, uncertainty could arise from an unknown existing infrastructure or unclear environmental and political factors. For example, the exact existing underground infrastructure might be unknown and potentially decrease the cost of building a connection, and the building of a pipeline could lead to conflicts with environmental protection groups or nearby residents that might increase the cost. These dynamic changes in the cost can be modeled by uncertainty intervals. Such uncertainties can then be resolved by inspecting the existing infrastructure or surveying residents and other potential stakeholder; both actions can be modeled by queries. Since the described actions to resolve the uncertainty can be cost extensive, the goal is to find an MST while minimizing the query cost.

It is well known that edges that have unique minimum weight in a cut of the graph are part of any MST. Furthermore, edges that have unique maximum weight on a cycle are part of no MST. Thus, to solve the MST problem under explorable uncertainty, we have to analyze the behavior of intervals and queries in terms of their interplay on cycles and in cuts. A simple cycle with three edges (triangle) already gives both lower bound examples and insights about the structure of a feasible query set. Consider the example of Fig. 3. It is clear that edge h is part of every MST, but we cannot decide which of the two edges f and g is in the MST without querying at least one of them. Similar to the lower bound example for the minimum problem, querying f first, in the worst case, reveals the green circles as precise weights, while querying g first reveals the red crosses. This forces any deterministic algorithm to query two elements, while the optimal query set contains just one. Thus, as was

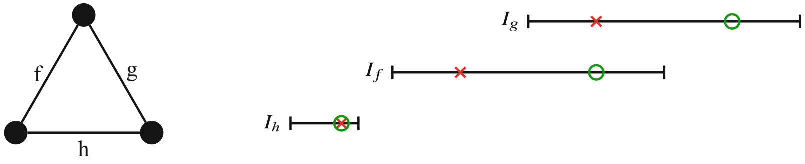


Fig. 3 Lower bound example for the minimum spanning tree problem

observed in Erlebach et al. (2008), no such algorithm can achieve a competitive ratio smaller than 2.

3.2 Mandatory Queries

A key aspect of several algorithms for problems under explorable uncertainty is the identification of *mandatory queries*. An interval $I_i \in \mathcal{I}$ is mandatory for the problem instance if each feasible query set has to query I_i, i.e., $I_i \in Q$ for all feasible query sets Q. The identification of mandatory queries is important since an algorithm can query such intervals without ever worsening its competitive ratio. In a sense, mandatory queries allow an algorithm to extract new information "for free." The revealed precise values might in turn allow the identification of further mandatory queries and, thus, lead to chains of mandatory queries. While it is possible to achieve theoretical worst-case guarantees without exploiting mandatory elements, empirical results indicate that the performance of algorithms significantly improves when the algorithm prioritizes the identification and querying of mandatory intervals (Erlebach et al. 2020; Focke et al. 2020).

When characterizing mandatory queries, we distinguish between characterizations based on the unknown precise values and characterizations that are only based on the uncertainty intervals. While the latter only uses information that can be accessed by an algorithm and, therefore, can actually be used to identify mandatory queries, the former is still helpful to analyze algorithms and will be useful in the following sections. We continue by characterizing mandatory queries for the two example problems.

3.2.1 Identifying Mandatory Queries for the Minimum Problem

Consider the minimum problem in multiple sets as introduced in the previous section. For a set $S \in \mathcal{S}$, we call an interval $I_i \in S$ the *precise minimum* of S if I_i has minimum precise value over all elements of S. The following lemma allows us to identify mandatory queries based on the precise values of the intervals.

Lemma 1 (Erlebach et al. (2020)) *An interval I_i is mandatory for the minimum problem if and only if (a) I_i is a precise minimum of a set S and contains w_j of another interval $I_j \in S \setminus \{I_i\}$ (in particular, if $I_j \subseteq I_i$), or (b) I_i is not a precise minimum of a set S with $I_i \in S$ but contains the value of the precise minimum of S.*

A common proof technique to show that an interval I_i is mandatory is to consider the query set $\mathcal{I} \setminus \{I_i\}$. Showing that querying every element except I_i does not solve the problem implies that I_i is mandatory. Vice versa, if querying $\mathcal{I} \setminus \{I_i\}$ solves the problem, then I_i is not mandatory. The following proof, which was given in Erlebach et al. (2020), uses this proof technique to show Lemma 1.

Proof If I_i is the precise minimum of S and contains w_j of another interval $I_j \in S$, then S cannot be solved even if we query all intervals in $S \setminus \{I_i\}$, as we cannot prove $w_i \leq w_j$ or $w_j \leq w_i$. If I_i is not a precise minimum of set S with $I_i \in S$ and contains the precise minimum value w^*, then S cannot be solved even if we query all intervals in $S \setminus \{I_i\}$, as we cannot prove that $w^* \leq w_i$.

If I_i is the precise minimum of a set S, but $w_j \notin I_i$ for every $I_j \in S \setminus \{I_i\}$, then $S \setminus \{I_i\}$ is a feasible query set for S. If I_i is not a precise minimum of a set S and does not contain the precise minimum value of S, then again $S \setminus \{I_i\}$ is a feasible query set for S. If every set S that contains I_i falls into one of these two cases, then querying all intervals except I_i is a feasible query set for the whole instance. □

Explicitly, Lemma 1 only enables us to identify mandatory intervals given full knowledge of the precise values, but it also implies criteria to identify *known mandatory* intervals, i.e., intervals that are known to be mandatory given only the intervals, and precise values revealed by previous queries. We call an interval *leftmost* in a set S if it is an interval with minimum lower limit in S. The following corollary follows from Lemma 1 and gives a characterization of known mandatory intervals.

Corollary 1 (Erlebach et al. (2020)) *If the leftmost interval I_l in a set S contains the precise value of another interval in S, then I_l is mandatory. In particular, if I_l is leftmost in S and $I_j \subseteq I_l$ for some $I_j \in S \setminus \{I_l\}$, then I_l is mandatory.*

3.2.2 Identifying Mandatory Queries for the Minimum Spanning Tree Problem

Mandatory queries for the MST problem can be characterized by using a structural property given by Megow et al. (2017). Let the *lower limit tree* $T_L \subseteq E$ be an MST for values w^L with $w_e^L = L_e + \epsilon$ for an infinitesimally small $\epsilon > 0$. Similarly, let the *upper limit tree* T_U be an MST for values w^U with $w_e^U = U_e - \epsilon$. Using the lower and upper limit trees, the following lemma allows us to identify mandatory queries based only on the intervals.

Lemma 2 (Megow et al. (2017)) *Any edge in $T_L \setminus T_U$ is mandatory.*

Thus, we may repeatedly query edges in $T_L \setminus T_U$ until $T_L = T_U$, and this will not worsen the competitive ratio. By this preprocessing, we may assume $T_L = T_U$. A characterization of the mandatory queries based on the full knowledge of the precise values is given by Erlebach and Hoffmann (2014).

3.3 Methods and Results

While the identification and querying of mandatory elements improve the performance of algorithms empirically and will be a key ingredient in the following sections, it is not sufficient to solve our two example problems. Therefore, we consider the *witness set algorithm*, one of the most important frameworks in explorable uncertainty. The witness set algorithm was introduced by Bruce et al. (2005) and relies on the identification of *witness sets*. A set $W \subseteq \mathcal{I}$ is a witness set if each feasible query set has to query at least one member of W, i.e., if $W \cap Q \neq \emptyset$ for all feasible query sets Q. Note that witness sets W with $|W| = 1$ are exactly the mandatory queries. Algorithm 1 formulates the witness set algorithm in a problem independent way. The algorithm essentially just queries witness sets until the problem is solved. Similar to mandatory queries, we distinguish between witness sets that can be identified based on the uncertainty intervals alone and witness sets that can only be identified based on knowledge of the precise values. The algorithm can only use the former kind.

Algorithm 1: Abstract formulation of the witness set algorithm

```
Input: Problem under explorable uncertainty with uncertainty intervals $\mathcal{I}$
while The problem is not solved yet do
    Query all elements of a witness set $W$.
```

The competitive ratio of the witness set algorithm depends on the size of the queried witness sets as formulated in the following lemma.

Lemma 3 (Bruce et al. (2005)) *If* $|W| \leq \rho$ *holds for all witness sets* W *that are queried by the witness set algorithm, then the algorithm is* ρ*-competitive.*

Proof Since querying elements multiple times does not reveal additional information, we can assume that all queried witness sets are pairwise disjoint. Let $W_1, \ldots, W_k$ denote those witness sets. Then, by definition of witness sets and since the sets are pairwise disjoints, the optimal query set contains at least k elements. By assumption, $|W_j| \leq \rho$ holds for all $j \in \{1, \ldots, k\}$. Thus, the algorithm queries at most $\rho \cdot k$ elements, and the competitive ratio is at most $\frac{\rho \cdot k}{k} = \rho$. □

In order to apply (and analyze) the witness set algorithm to a concrete problem, one has to characterize witness sets, bound the size of the witness sets, and show that the problem is solved once the characterization does not admit any more witness sets. In the following, we apply the algorithm to the two example problems.

3.3.1 Witness Set Algorithm for the Minimum Problem

For the minimum problem, we can identify witness sets of size one, i.e., mandatory queries, by using Corollary 1. Furthermore, we can identify witness sets of size two using the following lemma that was first (implicitly) shown by Kahan (1991).

Lemma 4 (Kahan (1991)) *A set* $\{I_i, I_j\} \subseteq \mathcal{I}$ *is a witness set if there exists an* $S \in \mathcal{S}$ *with* $\{I_i, I_j\} \subseteq S$, $I_i \cap I_j \neq \emptyset$, *and either* I_i *or* I_j *leftmost in* S.

Similar to the proof of the mandatory characterization, the lemma can be shown by considering the query set $Q = \mathcal{I} \setminus \{I_i, I_j\}$. After querying Q, both I_i and I_j still could be of minimum precise value in S. Thus, the problem is not solved yet, and at least one of I_i and I_j needs to be queried. This is a common proof strategy for showing that a subset of $\mathcal{I}$ is a witness set.

The witness set algorithm for the minimum problem repeatedly identifies and queries witness sets of size at most two by applying Corollary 1 and Lemma 4 until they cannot be applied anymore. If Lemma 4 cannot be applied anymore, then the leftmost interval I_i of each set S is not overlapped by any $I_j \in S \setminus \{I_i\}$. This implies that the leftmost intervals are the precise minima of the sets. Consequently, the problem then is solved, which implies the following theorem. The theorem was first (implicitly) shown by Kahan (1991) for a single set and translates to multiple sets.

Theorem 1 (Kahan (1991)) *The witness set algorithm is* 2*-competitive for the minimum problem. This competitive ratio is best possible for deterministic algorithms.*

3.3.2 Witness Set Algorithm for the Minimum Spanning Tree Problem

For the minimum spanning tree problem, we can identify witness sets of size one by using Lemma 2. Furthermore, we can identify witness sets of size two by using the following lemma that was shown in Erlebach et al. (2008), Megow et al. (2017). Recall that T_L and T_U are the lower and upper limit trees of the instance. Let $f_1, \ldots, f_l$ denote the edges in $E \setminus T_L$ ordered by non-decreasing lower limit, and let C_i be the unique cycle in $T_L \cup \{f_i\}$.

Lemma 5 (Erlebach et al. (2008)) *Let* $i \in \{1, \ldots, l\}$ *be the smallest index such that* $I_{f_i} \cap I_e \neq \emptyset$ *holds for some* $e \in C_i \setminus \{f_i\}$. *Then,* $\{f_i, e\}$ *is a witness set.*

The witness set algorithm for the MST problem repeatedly identifies and queries witness sets of size at most two by applying Lemmas 2 and 5 until they cannot be applied anymore. If Lemma 5 cannot be applied anymore, then each f_i does not overlap with any $e \in C_i \setminus \{f_i\}$. This implies that each f_i is maximal in C_i and therefore not part of any MST. Thus, T_L is known to be an MST and the problem is solved. This implies the following theorem.

Theorem 2 (Erlebach et al. (2008)) *The witness set algorithm is* 2*-competitive for the MST problem. This competitive ratio is best possible for deterministic algorithms.*

4 Explorable Uncertainty Beyond Worst-Case Analysis

In the previous section, we saw that, for both example problems, there is a 2-competitive algorithm with a matching lower bound. A natural question asks for ways to circumvent those lower bounds. There are different strategies for adjusting the model and performance guarantees beyond worst-case analysis. One common strategy is to allow an algorithm to make decisions randomly and measure the worst-case performance in expectation. The randomized algorithm given by Megow et al. (2017) shows that randomization is indeed powerful and admits improved results for the MST problem.

However, in this section, we follow a different approach and assume access to additional information on the problem instance. We present a *learning-augmented* and a *stochastic* variant of explorable uncertainty, where we are given predictions without any guarantee and probabilistic information, respectively. We outline how to design algorithms that can exploit this extra information by obtaining provable performance guarantees.

4.1 Exploiting Untrusted Predictions

In this section, we review the learning-augmented methods for explorable uncertainty that were introduced very recently by Erlebach et al. (2020). In the learning-augmented setting, we assume that we are given additional information on the problem instance in the form of *predictions* for the precise values of the uncertainty intervals. Those predictions could, for example, be derived by using machine learning (ML) methods. Based on the tremendous progress in artificial intelligence and machine learning, assuming access to such predictions of good accuracy seems reasonable. However, there is no guarantee on the accuracy and the predictions might be arbitrarily wrong. The learnability of predictions for the example problems is discussed in Erlebach et al. (2020).

Formally, we assume that we are given a predicted value $\overline{w}_i$ for each $I_i \in \mathcal{I}$. The predicted values are available *before* we query any elements and predict the results of the queries. Since we do not have any accuracy guarantee on the predictions, we call them *untrusted* to emphasize that the difference between w_i and $\overline{w}_i$ might be arbitrarily large. To compensate for the missing guarantee on the prediction quality, we aim at designing algorithms that achieve an improved performance for accurate predictions while being robust against arbitrarily bad predictions.

We refine competitive analysis to formulate those two objectives by adding a prediction awareness and adopt the notions of α-consistency and β-robustness (Lykouris and Vassilvitskii 2018; Purohit et al. 2018). An algorithm is α-consistent if it is α-competitive when the predictions are correct, i.e., $w_i = \overline{w}_i$ for all $I_i \in \mathcal{I}$, and it is β-robust if it is β-competitive for arbitrarily wrong predictions. While consistency and robustness only formulate the extreme cases for the prediction quality, we are also interested in guarantees with a smooth transition between consistency and robustness. We aim for performance guarantees that linearly degrade for an increased prediction error. This motivates interesting questions regarding suitable ways of measuring these errors.

4.1.1 Error Measures and Learnability

In the following, we consider the error measures introduced by Erlebach et al. (2020). A first simple and natural prediction error is the number of inaccurate predictions $k_\# = |\{I_i \in \mathcal{I} \mid w_i \neq \overline{w}_i\}|$. However, for the two example problems, a performance guarantee that, in terms of consistency, improves upon the lower bound of two and linearly degrades depending on $k_\#$ is not possible (Erlebach et al. 2020). The reason is that $k_\#$ completely ignores the structure of the intervals. (Similarly, using an error metric that depends on the distances between w_i and $\overline{w}_i$, e.g., $\sum_{I_i \in \mathcal{I}} |w_i - \overline{w}_i|$, would not be meaningful because only the order of the values and the interval end points matters for our problems.) To address this weakness, Erlebach et al. (2020) introduced two refined measures for the prediction quality.

Hop distance. Intuitively, for each interval $I_i \in \mathcal{I}$ with $w_i \neq \overline{w}_i$, this error measure counts the number h_i of lower and upper limits in $\mathcal{I} \setminus \{I_i\}$ that lie in the intervals $[w_i, \overline{w}_i)$ or $(w_i, \overline{w}_i]$. The hop distance of a given instance is then $k_h = \sum_{i=1}^{n} h_i$; see also the left part of Fig. 4. Note that the hop distance value h_i for a single interval I_i only depends on the *number* of interval borders that lie between w_i and $\overline{w}_i$ and, apart from that, is independent of the distance between w_i and $\overline{w}_i$. While the hop distance captures how the relations of the precise values to other intervals change compared to the predicted values, not every such change affects a feasible query set.

Mandatory query distance. To compensate for this fact, Erlebach et al. (2020) introduce another error measure based on the set $\mathcal{I}_R$ of mandatory intervals and the set $\mathcal{I}_P$ of prediction mandatory intervals, i.e., intervals that are mandatory under the assumption that all predictions are correct. The *mandatory query distance* is the size of the symmetric difference of $\mathcal{I}_P$ and $\mathcal{I}_R$, i.e., $k_M = |\mathcal{I}_P \Delta \mathcal{I}_R| = |(\mathcal{I}_P \cup \mathcal{I}_R) \setminus (\mathcal{I}_P \cap \mathcal{I}_R)| = |(\mathcal{I}_P \setminus \mathcal{I}_R) \cup (\mathcal{I}_R \setminus \mathcal{I}_P)|$. The right part of Fig. 4 shows an example for the mandatory query distance with $k_M = 1$. With respect to the precise values, both $\{I_1\}$ and $\{I_2, I_3, I_4\}$ are feasible query sets, and therefore, no interval is part of every feasible query set. This implies $\mathcal{I}_R = \emptyset$. Under the assumption that the

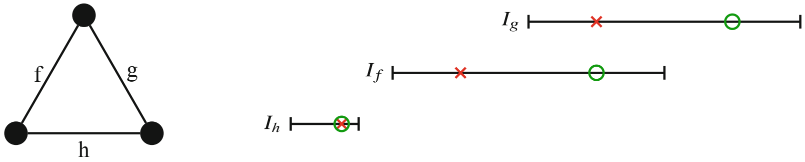


Fig. 4 Examples for the minimum problem with a single set $S = \{I_1, I_2, I_3, I_4\}$ as given in Erlebach et al. (2020). Circles and crosses illustrate precise values and predictions, respectively. Left: predictions and true values with a total hop distance of $k_h = 5$. Right: instance with a mandatory query distance of $k_M = 1$

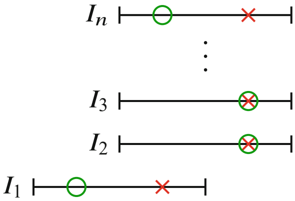


Fig. 5 Example for the minimum problem in a single set as given in Erlebach et al. (2020). The crosses and circles denote the predicted and precise values, respectively

predicted values are correct, Lemma 1 implies that I_1 is part of every feasible query set, and therefore, $\mathcal{I}_P = \{I_1\}$. It follows $k_M = |\mathcal{I}_P \Delta \mathcal{I}_R| = 1$.

In Erlebach et al. (2020), it is shown that, for both example problems, it is indeed possible to learn predictions with respect to k_h and k_M.

4.1.2 Methods and Results

A key aspect in the design of learning-augmented algorithms for explorable uncertainty is the identification of *prediction mandatory* intervals, i.e., intervals that are mandatory under the assumption that the predicted values are correct. For the two example problems, we can apply the characterizations of mandatory queries (cf. Lemma 1 and Erlebach and Hoffmann (2014)) that are based on the precise values under the assumption that the predicted values match the precise values. This allows us to identify all prediction mandatory queries. If we completely trust the predictions, we should just query all prediction mandatory elements. However, the example in Fig. 5 shows that completely trusting the predictions can lead to an arbitrarily bad robustness. In the example, the intervals $\{I_2, \ldots, I_n\}$ are prediction mandatory, but, if the predicted value of I_1 is wrong, it may be the case that the optimal query set contains only I_1.

This example implies that algorithms should balance the querying of prediction mandatory intervals with additional queries. In fact, Erlebach et al. (2020) show the following lower bound on the optimal trade-off between consistency and robustness for the two example problems.

Theorem 3 (Erlebach et al. (2020)) *Let $\beta \geq 2$ be a fixed integer. For the minimum and MST problems under explorable uncertainty, there is no deterministic β-robust algorithm that is α-consistent for $\alpha < 1 + \frac{1}{\beta}$. And vice versa, no deterministic α-consistent algorithm, with $\alpha > 1$, is β-robust for $\beta < \max\{\frac{1}{\alpha-1}, 2\}$.*

As a consequence of Theorem 3, no 2-robust algorithm can be better than 1.5-consistent. Thus, if we want to match the lower bound of 2 in terms of robustness, we should aim for 1.5-consistency. To achieve a good trade-off between consistency and robustness for the two example problems, we can use the two-phase framework as presented in Erlebach et al. (2020). Here, we describe on a high level the framework that can be used to achieve 1.5-consistency and 2-robustness for the two problems. The framework starts by querying all prediction mandatory elements. In a second stage, when no elements are prediction mandatory anymore, the algorithm has to decide which non-prediction mandatory elements to query. For the example problems, the algorithm has to decide which element of each witness pair it queries.

During the first phase, it is not sufficient to just query prediction mandatory elements since we already saw that this can lead to an arbitrarily bad robustness. Thus, each query to a prediction mandatory element is complemented with queries to further elements in such a way that they form witness sets. As the framework aims at breaking the lower bound of 2 in terms of consistency, it is not enough to form and query a size-2 witness set. To guarantee a consistency of 1.5, the framework instead identifies and queries sets of three elements for which we can guarantee that at least 2 of them must be contained in any feasible query set. Finding such elements based on the interval structure alone is not always possible as this would imply a 1.5-competitive algorithm for the non-learning augmented setting, which would contradict the lower bounds for the two example problems. Therefore, the framework identifies such sets under the assumption that the predictions are correct. After identifying such a set, its elements have to be queried adaptively, and, in case of wrong predictions, the algorithm has to compensate for the error by not querying all three elements. Querying all three elements in parallel might, in the case of wrong predictions, lead to a violation of the 2-robustness. The first framework phase repeatedly identifies such elements and queries them in a careful order while adjusting for potential errors, until no prediction mandatory elements remain. The characterization and identification of sets that satisfy the mentioned guarantee is problem specific and a challenging key aspect when applying this strategy to the concrete problems.

In the second phase of the framework, there are no more prediction mandatory elements, and the algorithm cannot identify any more "safe" queries. For the two example problems, this means that the algorithm has to decide, for each witness pair as characterized by Lemmas 4 and 5, which of the two elements to query. These decisions come down to finding a minimum vertex cover in an auxiliary graph representing the structure of the witness sets. In particular, the second phase for the minimum problem consists of finding and querying a minimum vertex cover. If the predictions are correct, querying the vertex cover solves the remaining problem. Otherwise, additional queries might be necessary, but those queries can be shown to be mandatory. Since both the size of the minimum vertex cover and the number of queried mandatory elements are lower bounds on the optimal query cost, the second framework phase is even 1-consistent and 2-robust. In the MST problem, wrong predictions can change the witness sets dynamically. Therefore, the second framework phase must be executed in a more adaptive and very careful way, requiring substantial additional work.

By applying (a generalized version of) the described framework, the following results for the two example problems can be achieved. In both theorems, OPT denotes the cost of an optimal query set, and the parameter γ can be used to configure the degree to which we trust the predictions. For increasing γ, the consistency improves but the robustness gets worse. To model smooth transitions between consistency and robustness, the theorems state the performance guarantees as the minimum of the error-dependent consistency and the robustness.

Theorem 4 (Erlebach et al. (2020)) *There is an algorithm for the minimum problem under uncertainty that, given an integer parameter* $\gamma \geq 2$*, achieves a competitive ratio of* $\min\{(1+\frac{1}{\gamma})(1+\frac{k_h}{OPT}), \gamma\}$*. If* $\gamma = 2$*, the algorithm is* 1.5*-consistent and* 2*-robust. Furthermore, there is an algorithm for the minimum problem under uncertainty that, given an integer parameter* $\gamma \geq 2$*, achieves a competitive ratio of* $\min\{(1+\frac{1}{\gamma-1}) \cdot (1+\frac{k_M}{OPT}), \gamma\}$.

Theorem 5 (Erlebach et al. (2020)) *There is a* 1.5*-consistent and* 2*-robust algorithm for the MST problem under uncertainty. Furthermore, there is an algorithm with competitive ratio* $\min\{1+\frac{1}{\gamma}+\left(5+\frac{1}{\gamma}\right)\cdot\frac{k_h}{OPT}, \max\{3, \gamma+\frac{1}{OPT}\}\}$*, for any* $\gamma \in \mathbb{Z}_{\geq 2}$.

These results show that learning augmentation can be successfully applied to problems under explorable uncertainty and circumvents known lower bounds for good predictions, while at the same time providing strong bounds on the worst-case performance even when the predictions are completely wrong. This eases the integration of machine learning into a system since it allows improved results while protecting users from occasional failures of the ML algorithms.

4.2 Exploiting Stochastic Information

Recently, the setting of *stochastic explorable uncertainty* has received some attention in the context of sorting (Chaplick et al. 2020) and the minimum problem (Bampis et al. 2021), which Bampis et al. (2021) phrases as a (hyper-)graph orientation problem.

In the stochastic setting, we are given a continuous probability distribution d_i over the interval $I_i = (L_i, U_i)$ for each $I_i \in \mathcal{I}$. The precise value w_i of an interval I_i

is drawn independently from d_i. Bampis et al. (2021) again analyze an algorithm ALG in terms of its competitive ratio

$$\max_{J \in \mathcal{J}} \frac{\mathbb{E}[\text{ALG}(J)]}{\mathbb{E}[\text{OPT}(J)]},$$

where $\mathbb{E}[\text{ALG}(J)]$ denotes the expected query cost of ALG when solving instance J, $\mathbb{E}[\text{OPT}(J)]$ denotes the expected optimal query cost for J, and $\mathcal{J}$ is the set of all instances. As a main result, Bampis et al. (2021) give the following theorem.

Theorem 6 (Bampis et al. (2021)) *For any $\epsilon > 0$, there exists a $f(\alpha)$-competitive algorithm for the minimum problem in multiple sets, where $f(\alpha) \in [1.618 + \epsilon, 2]$ depends on the approximation ratio α for solving a vertex cover problem in an auxiliary graph.*

The algorithm relies on computing the probability that an interval is mandatory using the characterization of Lemma 1. It queries all vertices that have a mandatory probability exceeding a certain threshold. Afterward, the algorithm solves a vertex cover problem on an auxiliary graph by first preprocessing the instance via a classical linear programming relaxation and, afterward, executing the α-approximation. The algorithm queries the computed vertex cover and, thereafter, only mandatory intervals that remain. In addition to this general algorithm, Bampis et al. (2021) give several lower bounds and improved algorithms for special cases that also prioritize queries to intervals with a high probability to be mandatory.

5 Concluding Remarks

This chapter discusses the model of explorable uncertainty and its potential use for decision-making under uncertainty in logistics. We illustrate classical techniques to design algorithms with worst-case guarantees using the two example problems of finding the minima of multiple sets and determining a minimum spanning tree. With the learning-augmented and stochastic setting, we also present models and techniques for algorithm design with guarantees beyond the worst case and show that known limitations of worst-case analysis can be circumvented by such settings.

In this chapter, we require algorithms to determine an optimal solution for the underlying optimization problem. One could relax this restriction and ask for an α-approximation. Unfortunately, for the example problems, the lower bounds translate to this relaxed setting (Megow et al. 2017, Section 10). Another interesting variation of the model refers to the objective function. While we consider settings in which the query cost is significant, optimization with explorable uncertainty seems relevant also in settings where query cost and objective value of the underlying problem are comparable. In this case, it would be interesting to consider a combined objective, e.g., to minimize the (weighted) sum of both values.

We illustrated techniques using two example problems, a selection and a network design problem. These appear as sub-problems in classical logistics questions and our techniques would be directly applicable. Admittedly, a rigorous worst-case guarantee such as the competitive ratio may be relevant only in rare applications; otherwise, a good empirical performance on practical input instances is sufficient. In more complex problem settings, e.g., involving additionally routing, packing, and scheduling aspects, we may not be able to prove worst-case guarantees on the performance of our algorithms and they may not even exist. Here, performance measures beyond the worst case are particularly relevant. Overall, we expect that the model and techniques presented here give insights on the power, tractability, and applicability of explorable uncertainty, and we hope that they can serve as a first step toward tackling more complex logistics problems.

6 Bibliographical Notes

We conclude with further pointers to previous work on optimization with explorable uncertainty.

The line of research on explorable uncertainty has been initiated by Kahan (1991) in the context of selection problems. Subsequent work addressed caching problems (Olston and Widom 2000), problems such as computing a function value (Khanna and Tan 2001), finding the kth smallest value in a set of uncertainty intervals (Feder et al. 2003; Gupta et al. 2016), also with non-uniform query cost (Feder et al. 2003), and sorting (Halldórsson and de Lima 2019).

Interestingly, the sorting problem is a special case of the minimum problem for multiple sets. Given an instance of the sorting problem, we can create a set for each pair of elements that are in the same set of the sorting instance and obtain a minimum problem whose feasible query sets also solve the sorting problem. Halldórsson and de Lima (2019) showed directly that the witness set algorithm for sorting a single set is 2-competitive and is best possible. They also show that the competitive ratio can be improved to 1.5 using randomization. Furthermore, Chaplick et al. (2020) introduce an algorithm for sorting a single set of elements under stochastic uncertainty that is optimal in terms of the expected cost $\mathbb{E}[\textsc{alg}(J)]$. The competitive ratio of this algorithm is unknown.

Only more recently also optimization problems have been studied. A key role plays the fundamental MST problem with uncertainty. The 2-competitive deterministic witness set algorithm was presented and shown to be best possible by Erlebach et al. (2008). The randomized algorithm by Megow et al. (2017) has an improved competitive ratio of 1.707. Both a deterministic 2-competitive algorithm and a randomized 1.707-competitive algorithm are known for the more general problem of finding the minimum base in a matroid (Erlebach et al. 2016; Megow et al. 2017), even for the case with non-uniform query costs (Megow et al. 2017). Other works on the MST problem (and matroids) study a non-adaptive variant (Merino and Soto 2019) and the offline verification problem (Erlebach and

Hoffmann 2014; Megow et al. 2017) and conduct an experimental study (Focke et al. 2020).

Other optimization problems studied in the context of explorable uncertainty include the shortest path problem (Feder et al. 2007), the knapsack problem (Goerigk et al. 2015), and scheduling problems (Albers and Eckl 2020; Arantes et al. 2018; Dürr et al. 2020).

The growth of data-driven applications and machine learning methods in the past years gave rise to a model for learning-augmented online algorithms. The model has been proposed by Medina and Vassilvitskii (2017) in the context of revenue optimization followed by work on online caching by Lykouris and Vassilvitskii (2018). Purohit et al. (2018) studied online scheduling and rent-or-buy problems with respect to consistency and robustness, and they obtained performance guarantees as a function of the prediction error. This work initiated a vast growing line of research, which makes ML predictions without any accuracy guarantee useful in the design of algorithms with hard performance guarantees. Overall, learning-augmented online optimization is a highly topical concept with high potential also for applications in logistics problems.

Acknowledgments The authors were partially funded by the German Science Foundation (DFG).

References

Albers, S., Eckl, A.: Explorable uncertainty in scheduling with non-uniform testing times. In: International Workshop on Approximation and Online Algorithms (2020)

Arantes, L., Bampis, E., Kononov, A.V., Letsios, M., Lucarelli, G., Sens, P.: Scheduling under uncertainty: a query-based approach. In: IJCAI 2018: 27th International Joint Conference on Artificial Intelligence, pp. 4646–4652 (2018)

Bampis, E., Dürr, C., Erlebach, T., de Lima, M.S., Megow, N., Schlöter, J.: Orienting (hyper)graphs under explorable stochastic uncertainty. In: Proceedings of the 29th Annual European Symposium on Algorithms (2021)

Ben-Tal, A., El Ghaoui, L., Nemirovski, A.S.: Robust Optimization. Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2009)

Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming, 2nd edn. Springer, New York (2011)

Borodin, A., El-Yaniv, R.: Online Computation and Competitive Analysis. Cambridge University Press, Cambridge (1998)

Bruce, R., Hoffmann, M., Krizanc, D., Raman, R.: Efficient update strategies for geometric computing with uncertainty. Theory Comput. Syst. **38**(4), 411–423 (2005)

Chaplick, S., Halldórsson, M.M., de Lima, M.S., Tonoyan, T.: Query minimization under stochastic uncertainty. In: Kohayakawa, Y., Miyazawa, F.K. (eds.) LATIN 2020, vol. 12118. Lecture Notes in Computer Science, pp. 181–193 (2020)

Dürr, C., Erlebach, T., Megow, N., Meißner, J.: An adversarial model for scheduling with testing. Algorithmica **82**(12), 3630–3675 (2020)

Erlebach, T., Hoffmann, M.: Minimum spanning tree verification under uncertainty. In: Kratsch, D., Todinca, I. (eds.) International Workshop on Graph-Theoretic Concepts in Computer Science, vol. 8747. Lecture Notes in Computer Science, pp. 164–175 (2014)

Erlebach, T., Hoffmann, M., Krizanc, D., M. Mihalák, Raman, R.: Computing minimum spanning trees with uncertainty. In: Proceedings of the 25th Annual Symposium on the Theoretical Aspects of Computer Science, pp. 277–288 (2008)

Erlebach, T., Hoffmann, M., Kammer, F.: Query-competitive algorithms for cheapest set problems under uncertainty. Theor. Comput. Sci. **613**, 51–64 (2016)

Erlebach, T., Hoffmann, M., de Lima, M.S., Megow, N., Schlöter, J.: Untrusted predictions improve trustable query policies. CoRR, abs/2011.07385 (2020)

Feder, T., Motwani, R., Panigrahy, R., Olston, C., Widom, J.: Computing the median with uncertainty. SIAM J. Comput. **32**(2), 538–547 (2003)

Feder, T., Motwani, R., O'Callaghan, L., Olston, C., Panigrahy, R.: Computing shortest paths with uncertainty. J. Algorithms **62**(1), 1–18 (2007)

Focke, J., Megow, N., Meißner, J.: Minimum spanning tree under explorable uncertainty in theory and experiments. ACM J. Exp. Algorithmics **25**, 1–20 (2020)

Goerigk, M., Gupta, M., Ide, J., Schöbel, A., Sen, S.: The robust knapsack problem with queries. Comput. Operat. Res. **55**, 12–22 (2015)

Gupta, M., Sabharwal, Y., Sen, S.: The update complexity of selection and related problems. Theory Comput. Syst. **59**(1), 112–132 (2016)

Halldórsson, M.M., de Lima, M.S.: Query-competitive sorting with uncertainty. In: 44th International Symposium on Mathematical Foundations of Computer Science, vol. 138. Leibniz International Proceedings in Informatics, pp. 7:1–7:15 (2019)

Kahan, S.: A model for data in motion. In: STOC'91: 23rd Annual ACM Symposium on Theory of Computing, pp. 265–277 (1991)

Khanna, S., Tan, W.-C.: On computing functions with uncertainty. In: PODS '01: Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, pp. 171–182 (2001)

Lang, W., Jedermann, R.: What can MEMS do for logistics of food? Intelligent container technologies—a review. IEEE Sensors J. **16**(18), 6810–6818 (2016)

Lykouris, T., Vassilvitskii, S.: Competitive caching with machine learned advice. In: Proceedings of Machine Learning Research, pp. 3302–3311 (2018)

Medina, A.M., Vassilvitskii, S.: Revenue optimization with approximate bid predictions. In: Proceedings of NIPS, pp. 1856–1864 (2017)

Megow, N., Meißner, J., Skutella, M.: Randomization helps computing a minimum spanning tree under uncertainty. SIAM J. Comput. **46**(4), 1217–1240 (2017)

Merino, A.I., Soto, J.A.: The minimum cost query problem on matroids with uncertainty areas. In: Proceedings of ICALP, vol. 132. Leibniz International Proceedings in Informatics, pp. 83:1–83:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Wadern (2019)

Olston, C., Widom, J.: Offering a precision-performance tradeoff for aggregation queries over replicated data. In: VLDB 2000: 26th International Conference on Very Large Data Bases, pp. 144–155 (2000)

Purohit, M., Svitkina, Z., Kumar, R.: Improving online algorithms via ML predictions. In: Proceedings of Advances in Neural Information Processing Systems, pp. 9661–9670 (2018)

Sanchez-Rodrigues, V., Potter, A., Naim, M.M.: Evaluating the causes of uncertainty in logistics operations. Int. J. Logist. Manag. **21**(1), 45–64 (2010)

Simangunsong, E., Hendry, L.C., Stevenson, M.: Supply-chain uncertainty: a review and theoretical foundation for future research. Int. J. Prod. Res. **50**(16), 4493–4523 (2012)

van der Vorst, J.G., Beulens, A.J.: Identifying sources of uncertainty to generate supply chain redesign strategies. Int. J. Phys. Distrib. Logist. Manag. **32**(6), 409–430 (2002)

Wilding, R.: The supply chain complexity triangle: uncertainty generation in the supply chain. Int. J. Phys. Distrib. Logist. Manag. **28**(8), 599–616 (1998)

Complex Networks in Manufacturing and Logistics: A Retrospect

Till Becker and Darja Wagner-Kampik

Abstract The methodology to model systems as graphs or networks already exists for a long time. The availability of information technology and computational power has led to a renaissance of the network modeling approach. Scientists have collected data and started to create huge models of complex networks from various domains. Manufacturing and logistics benefits from this development, because material flow systems are predetermined to be modeled as networks. This chapter revisits selected advances in network modeling and analysis in manufacturing and logistics that have been achieved in the last decade. It presents the basic modeling concept, the transition from static to dynamic and stochastic models, and a collection of examples how network models can be applied to contribute to solving problems in planning and control of logistic systems.

1 Introduction

Since the beginning of the division of labor, manufacturing processes have been split up in different tasks. These individual tasks have their own tools, materials, and places. To exploit the advantages of this type of organization of a manufacturing process, a proper planning and control of all involved aspects is vital. This chapter focuses on a manufacturing system's underlying structure, i.e., the pattern of the

---

Till Becker was affiliated to LogDynamics until 2018 as Head of the Research Group Production Systems and Logistics Systems at the University Bremen, Germany

---

T. Becker (✉)
Faculty of Business Studies, University of Applied Sciences Emden/Leer, Emden, Germany
e-mail: till.becker@hs-emden-leer.de

D. Wagner-Kampik
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany
e-mail: wan@biba.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_3

material flow. Such a pattern can be found on various levels of detail. In all cases, it is a network of entities connected by material flow. This can be, e.g., a network of machines on a shop floor or a network of suppliers, manufacturers, and retailers in a supply chain.

We present a retrospect on the development of a deeper understanding of complex networks in manufacturing and logistics throughout the last decade around the work of the *Production Systems and Logistic Systems* group within the *Research Cluster for Dynamics in Logistics* at the University of Bremen, Germany.

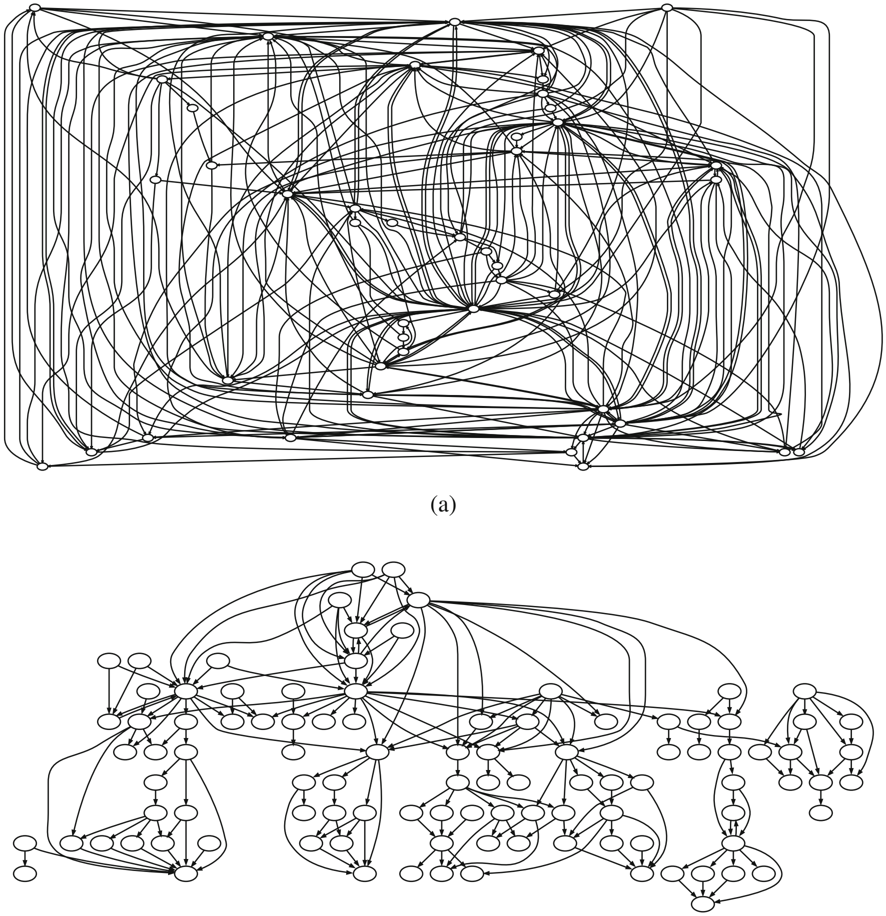
2 Complex Networks in Manufacturing and Logistics

2.1 Modeling of Complex Networks

Systems of interacting or connected entities can be modeled as a graph (G), which consists of a set of vertices (V) connected by a set of edges (E), such that $G = (V, E)$. Another widely used nomenclature calls these system networks, composed of nodes connected by links. In general, there is no difference in the meaning of the two variants and they can be used synonymously. Although the concept of graphs is known in mathematics since centuries, the modeling of large-scale systems as graphs has not been feasible for a long time due to the required effort to manually gather data, draw the graph, and perform any type of calculation for graph-based algorithms.

The pioneering research in complex networks has covered network models from various domains, such as social networks, biological networks, and computer networks (see, e.g., Albert and Barabási 2002; Boccaletti et al. 2006). The researchers have investigated the structure of complex networks, their properties, and their formation. A main finding throughout the years was the fact that network models, although coming from different domains, share common properties with regard to their structure.

The transfer and application of complex network theory to the manufacturing and logistics domain began a decade ago with the works of Vrabič et al. (2013) and Becker et al. (2014). The early contributions demonstrate how the material flow in a manufacturing system can be converted into a network model. Network modeling is based on material flow data, which can be collected as a byproduct from shop floor control systems. Such data consists of a list of all available machines on the shop floor and a collection of records documenting the individual material flow events. Each record contains at least the manufacturing order ID, the machine that processed a certain operation, as well as a timestamp. An algorithm can be used to build the network model by creating the sets V and E (Becker et al. 2014). The network nodes are represented by the machines on the shop floor. The set of edges is filled by linking each record with its consecutive record from the same manufacturing order: if an item from the same manufacturing order is first being



(a)

(b)

Fig. 1 Two network representations of the material for two different manufacturing scenarios created with data from manufacturing execution systems. (**a**) Network model of a shop floor production of machine parts. (**b**) Network model from process industry

processed on machine A, and the next timestamp is an operation on machine B, then a link $A \rightarrow B$ is added to E.

Figure 1 displays two network models created by Becker et al. (2014). Both were created in the exactly same fashion using the algorithm presented above. However, due to the differences in the handling of the material flow in the two systems, two very distinct network models were obtained. Figure 1a shows the material movements on a shop floor during the manufacturing of machine parts, whereas Fig. 1b illustrates the flow of items in process industry with distinct groups of machines in clusters.

2.2 *The Structure of Manufacturing Networks and Its Impact on Material Flow*

2.2.1 Comparison of Manufacturing Networks to Other Flow-Oriented Networks

A large part of the research work dealing with complex networks in manufacturing either discovers structural characteristics of those networks or investigates the relation between structural network characteristics and the performance of the material flow. Network models from other domains have already been described in detail, such as engineered networks from communication (Braha and Bar-Yam 2006), urban traffic (Lämmer et al. 2006), or supply chains (Meepetchdee and Shah 2007), but also evolved network structures like river networks (De Menezes and Barabási 2004) or predator–prey relations in ecosystems (Williams et al. 2002).

Therefore, a first approach was to compare the properties of manufacturing networks and other flow-oriented networks. The comparisons covered traffic networks (Becker et al. 2011) as well as metabolic networks (Becker et al. xxxx, 2011). These network types were chosen for comparison, because they also represent a structure in which a flow of items needs to be managed in order to fulfill the system's objectives. Networks of metabolic reactions in cells in particular can be seen as "factories," which transform incoming items into final products via enzyme reactions. The main findings were that, on the one hand, all these network types exhibit strong structural similarities when observing the distribution of network characteristics, such as degree of connectivity. On the other hand, flow simulations on these networks showed a distinct behavior. Metabolic networks tend to have a more evenly distributed flow, whereas manufacturing networks tend to go toward a hub-and-spoke architecture with a small number of highly frequented nodes.

2.2.2 The Relation Between Structure and Performance

As a consequence of the investigation of different networks from various domains, researchers have hypothesized that there is a relation between the static structure of a network and its performance regarding the material flow. The rationale behind this is the idea that networks are designed (or in the case of biological or ecological networks have evolved) in a way to support the material flow processes. In order to quantify structural properties of networks, a variety of network figures has been developed (see, e.g., Boccaletti et al. 2006). Most researchers focus on centrality measures, which indicate how strong the nodes in a network are connected with each other. The most frequently used centrality measures are Degree-Centrality, Betweenness-Centrality, Closeness-Centrality, and Eigenvector-Centrality (Becker and Wagner 2016). Centrality measures can be determined for individual nodes or as an average over all nodes in the network (Becker et al. 2012; Omar et al. 2018). A key observation by Becker et al. (2012) is the fact that there is a

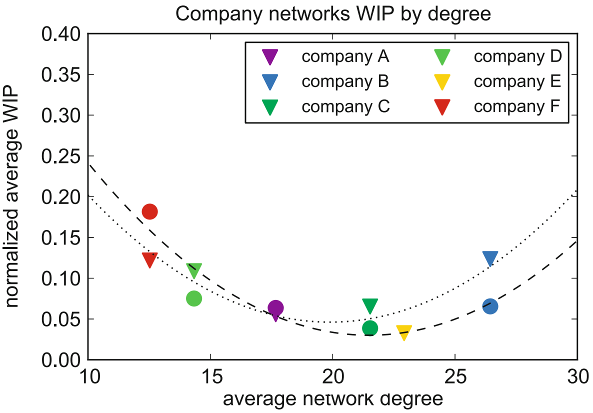


Fig. 2 The graph shows the relation between the average degree of a network and the work in process obtained from a series of material flow simulations. The results indicate that a specific degree of connectivity in a material flow network is best suited to prevent the buildup of queues (Becker et al. 2012)

nonlinear relation between the degree of connectivity in a material flow network and its performance. Figure 2 shows the results from a simulation study. A material flow simulation was carried out in a selection of networks. These networks were constructed from real-world data following the procedure described in Sect. 2.1. The *work in process* (WIP) served as key performance figure. It indicates how much work content is present in the system, e.g., waiting in queues or being currently processed, and should be kept on a low level. The obtained results show that there is an ideal degree of connectivity in the network structure which allows to keep the WIP low. Lower connectivity and increased connectivity lead to an increase in WIP.

Beber and Becker (2014) extended the scope beyond individual nodes and investigate patterns in the shape of three-node subgraphs, also known as network motifs. Their findings indicate that the motif signature can serve as a "fingerprint" to distinguish different manufacturing networks from each other.

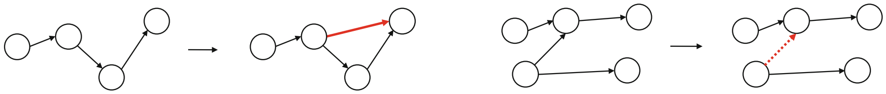
Other approaches make use of network centrality metrics to identify bottleneck work stations in a manufacturing system (Blunck et al. 2014). Although the assessment of the actual material flow in a network sill provides more reliable information whether a work station is a bottleneck or not, the centrality of a node in a network can also be used to identify those bottlenecks. The advantage of the network measures is their availability. Even in early planning stages, when material flow data is either not yet available or can only be acquired with the help of computer simulations, the bottleneck situation in a manufacturing system can already be assessed using network measures.

Another application-oriented approach using centrality measures is able to identify key machines in a manufacturing system based on network measures (Becker and Wagner 2016). A selection of centrality measures were used to identify the machines that have the highest impact on material flow performance. The results emphasized the applicability of certain network centrality measures for this purpose. Moreover, it could be shown that not all centrality measures are suited to identify important nodes in manufacturing systems, due to the specific mechanics of flow in these systems in comparison to other networks.

2.3 Dynamic Processes on Material Flow Networks

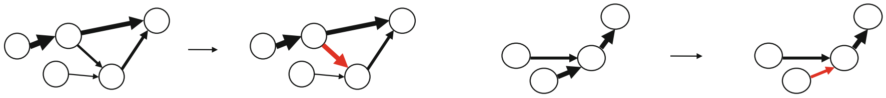
The previous modeling approach of networks basically considers the static description of material flow and information flow structures, which effectively means that network elements do not change over time. Applied to material flow systems, it would mean there is no change in material flow level at any time. However, recent studies suggest that the structure of a material flow network changes as a result of events or changing circumstances (e.g., passing a job to the next work station or failure of work stations) over a specific period of time (see, e.g., Beber and Becker 2014; Vrabič et al. 2013). Consequently, it might be useful to consider such dynamic processes for specific applications.

Figure 3 shows potential structural changes in networks. As indicated earlier, all these changes are, among other things, triggered by additional work systems, machine breakdowns, changes in product mix, etc. In real material flow systems, a combination of all these events leads to a variety structural changes during a certain observation period. For analytical purposes, the entire observation period can be divided into a number of time windows. If there are any active nodes and edges within a defined time window, a network instance is created. This way, structural changes can be observed and linked to various events. Although the length of each time window can be selected freely, it has a significant impact on the structural changes. For example, for shorter time windows, the average values of different key network figures display stronger variation from one time window to



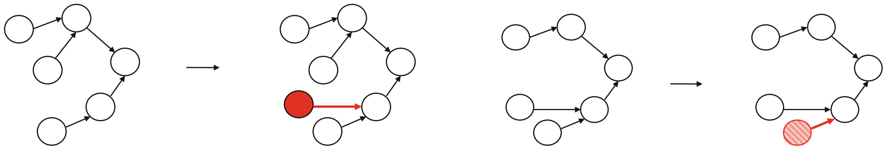
new material flow

disappearing material flow



intensified material flow

reduced material flow



new work system

disappearing work system

Fig. 3 The structural changes affect in particular both network elements (i.e., nodes and edges). It is therefore possible that nodes and edges disappear or join the network over time. Moreover, the intensity of the edge weights may vary over time (Wagner and Becker 2016) another, whereas for longer time windows the deviations between time windows are lower (Wagner and Becker 2016).

3 Advanced Network Modeling: Stochastic Block Models

Network modeling in manufacturing and logistics has seen a number of advances in the recent decade. The first simple, yet effective modeling approach based on material flow data has been presented in Sect. 2.1. The consideration of the development of networks over time was discussed in Sect. 2.3. A further development in network modeling in manufacturing was the introduction of stochastic models. In particular, the Stochastic Block Model (SBM) allows for a prediction of future states in a manufacturing system (Funke and Becker 2020). The motivation behind the application of SBMs as a tool in manufacturing systems modeling is the fact that many material flow systems consist of elements that can be grouped into clusters of similar objects in terms of their role in the material flow, e.g., manufacturing cells. In a nutshell, an SBM is a network model in which groups of similar nodes (like clusters) are seen as structural equivalent. Instead of explicitly modeling the links between nodes, the general probability of two nodes from two groups being connected is given (see Fig. 4).

Creating an SBM is not as straightforward as creating a simple network model as presented in Sect. 2.1. The modeling requires the selection of the desired type of SBM and an appropriate inference method to derive the actual model from the material flow data. Funke and Becker (2019) have investigated, compared, and evaluated a variety of SBM variants and inference methods to facilitate the selection process.

Funke and Becker (2020) were then able to demonstrate how an SBM can be applied to perform link prediction. Due to the stochastic nature of the model, it is possible to retrieve a probability of two nodes being connected (or not being connected) in the future. Consequently, managers and planners will be enabled to make design decision regarding the planning and control of material flow in their manufacturing systems.

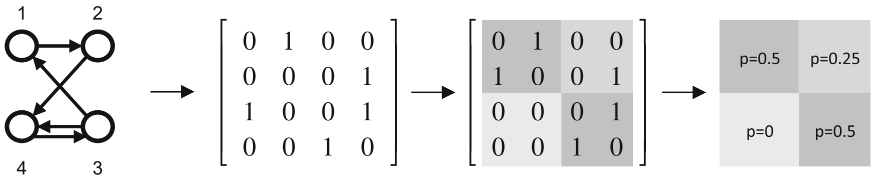


Fig. 4 Minimal model of an SBM: (1) the network model is transformed into an adjacency matrix, (2) the nodes are grouped into blocks with similar connections, and (3) the result is a block matrix with probabilities indicating the likelihood of two nodes from corresponding blocks being connected

4 Identification of Autonomous Clusters Considering the Topological Setting

Nowadays, companies are faced with constantly increasing complexity due, for example, to shorter product life cycles, increasing customer requirements, and high fluctuations in demand. This requires novel control solutions. The massive technological improvements over the past years additionally support the transformation toward autonomous control approaches. Previous research activities in the context of autonomous control are primarily limited to control algorithms and technologies. However, most companies are facing the challenge not only to select appropriate control algorithms and to adapt the new technologies but also to decide how the structure of a system, i.e., the network topology, can support their transformation toward decentralized autonomous control. From a topological perspective, modularization, which subdivides a system into small units with autonomous decision-making, seems to be a promising approach (Gronau and Theuer 2011; Mourtzis and Doukas 2013). The resulting question is which work systems should form a unit and act autonomously?

A first attempt to identify highly interacting work systems, in order to be able to subsequently merge them into autonomous units—here called clusters, was made by Vrabič et al. (2012). Their approach is based on the assumption that edges in networks represent the intensity of material flow between the corresponding nodes. Consequently, nodes inside a cluster should be strongly connected, but only loosely connected to other nodes outside the cluster. In recent years, a number of clustering methods have been developed. Some methods have been tested to determine their suitability for the identification of autonomous clusters. Subsequently, the proposed approach was extended to include the dynamic development of the manufacturing system over time (Becker and Wagner 2015; Becker and Weimer 2014). To this end, they evaluated the consistency of clusters over time (for an example, see Fig. 5).

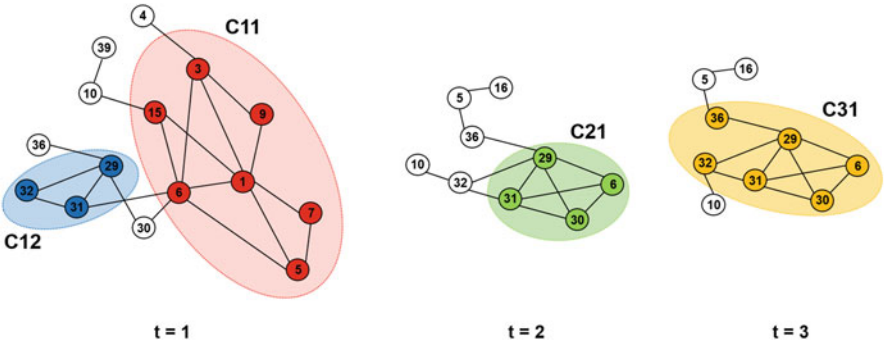


Fig. 5 Resulting network representations for three different time windows in order to reveal similar cluster pattern. Within the first time window, two clusters have been identified and only one cluster in the following time windows. The nodes 6, 29, 30, and 31 are represented in all time windows (Becker and Wagner 2015)

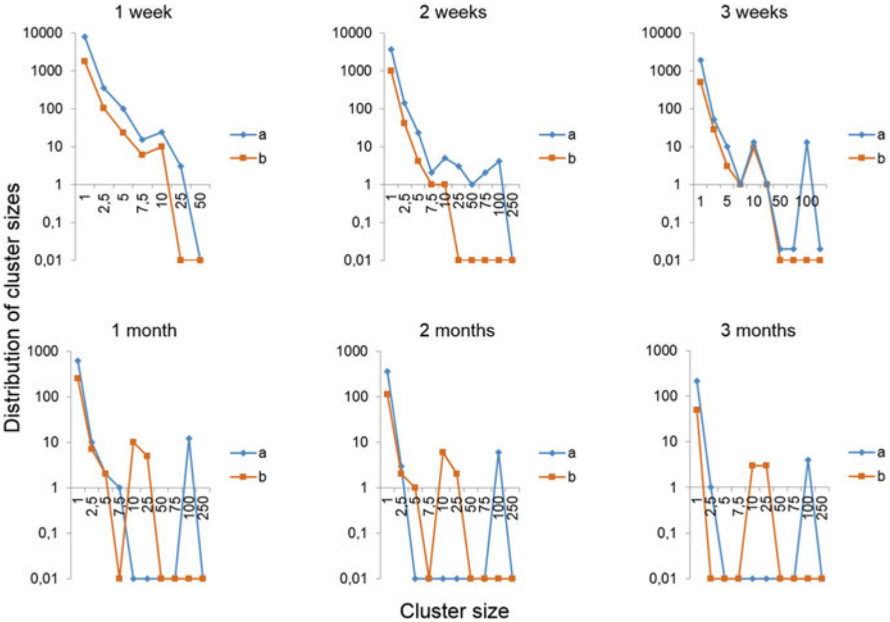


Fig. 6 Increasing the length of time windows results in an increasing number of nodes and edges. As a result, this leads to a lower number of clusters, which are in turn larger in size. The results for shorter time windows are mostly similar for both datasets (a and b)

Tracking the evolution of each cluster can help identify robust clusters, which are therefore particularly suited to serve as autonomous clusters. Furthermore, in accordance with the results of Lancichinetti et al. (2010) that real-world networks of the same category are identical with respect to their cluster structure, Wagner and Becker (2016) were able to show that this is also true for material flow networks for relatively shorter time windows (see Fig. 6).

5 Synthetic Material Flow Networks with a Built-In Cluster Structure: A Random Walk-Based Approach

In order to be able to test the proposed clustering approach for the identification of autonomous structures in networks of different typologies, appropriate data is needed. However, due to a lack of reliable data with different cluster structures, the necessity arises to generate well-characterized synthetic material flow networks that allow extensive testing in simulation studies. Synthetic networks are characterized by the fact that they are generated randomly, but within a framework of rules. To generate graphs with a certain built-in cluster structure, a random walk-based approach can be applied, as presented in Wagner and Becker (2017, 2018). With this approach, different job routing patterns are modeled as random walks.

A random walk is considered as a stochastic process which results in a sequence of randomly visited nodes (Lovász 1993). Consequently, random walks do not require any global network information. Moreover, they can help to understand the underlying routing mechanics of systems and are therefore able to reveal specific characteristics like cluster structures in networks (Schaeffer 2007). This is explained by the fact that random walks tend to get trapped in clusters. Hence, it is not the random walks themselves that create networks with certain characteristics, but they need a network as a basis for their movements. Such networks are also referred to as underlying networks (Rosvall and Bergstrom 2008). The topological characteristics of such a network have been used to guide a random walk. To be able to create networks with various built-in cluster structures, not only the parameters of the random walks have to be defined but also those of the underlying network.

After applying this approach, it can be observed how random walks change the number of clusters, the cluster sizes, and the degree of connectivity of clusters given in the underlying network. The main result shows that the random walk-based routing is well suited to map the cluster structures of a number of tested underlying networks (Wagner and Becker 2018). This gives the underlying network a leading function. In future, this approach can be extended to create networks with other topological patterns.

6 Summary and Outlook

Research on complex networks has become visible across many different disciplines throughout the recent years. With regard to the flow of material, goods, and information in supply chains, traffic, and on shop floors, network models are a powerful tool in manufacturing and logistics research. This chapter has shown the foundations of network modeling in this domain and presented a selection of more detailed research activities and application-oriented approaches. It became apparent that, on the one hand, many network-related insights from physics, biology, or social sciences can be transferred to manufacturing and engineering. On the other hand, material flow in manufacturing has also specific properties, which are distinct from the flow of traffic on roads or the flow of messages in social networks. This is why it is essential to continue the research on complex networks in manufacturing and logistics.

The direction for future research should be aligned with the application opportunities in industry. More and more companies will collect massive amounts of system data in the context of Industry 4.0 and thus will be able to make use of network modeling. The inclusion of Artificial Intelligence (AI) methods in the processing of network data can further increase the quality of network models. Therefore, the so far developed methodologies should be transferred into practical applications for decision-making in planning, operation, and control of logistic systems. Figure 7 summarizes the development of network research in manufacturing and logistics.

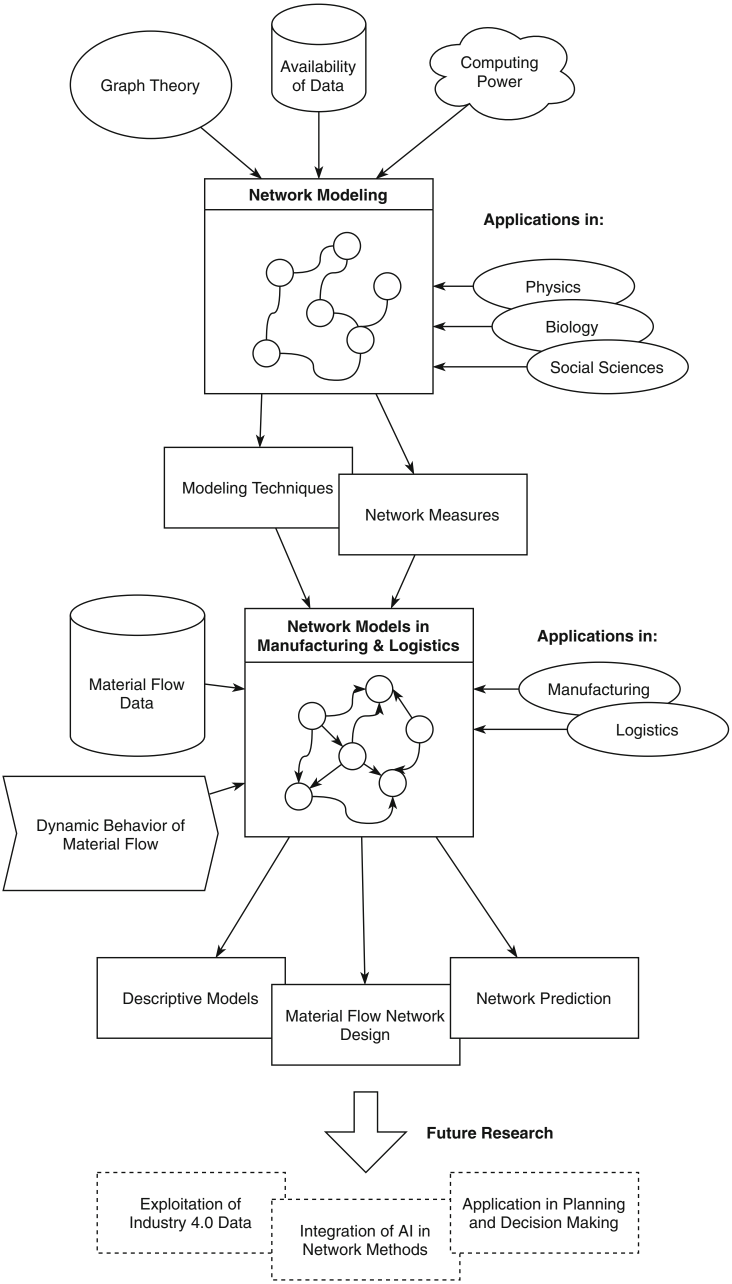


Fig. 7 Network modeling in manufacturing and logistics has emerged from the application of graph theory in the information age. Initially, researchers have concentrated on modeling and description of logistic networks. Over the years, the research has evolved into a multitude of methods for description, design, and prediction of networks. Future research should be focused on creating industrial applications and on the inclusion of additional data sources and AI methods

Acknowledgments The authors and the presented work were partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grant numbers BE 5538/2-1 and BE 5538/3-1, as well as by the Institutional Strategy of the University of Bremen, and funded by the German Excellence Initiative.

References

Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002). https://doi.org/10.1103/RevModPhys.74.47

Beber, M.E., Becker, T.: Towards an understanding of the relation between topological characteristics and dynamic behavior in manufacturing networks. In: Procedia CIRP, vol. 19. 2nd CIRP Robust Manufacturing Conference (RoMac 2014), pp. 21–26 (2014). https://doi.org/10.1016/j.procir.2014.05.005

Becker, T., Wagner, D.: Stufenweise Einführung von dezentraler Steuerung. Ind. Manag. **31**(2), 28–32 (2015)

Becker, T., Wagner, D.: Identification of key machines in complex production networks. In: Procedia CIRP, vol. 41. Research and Innovation in Manufacturing: Key Enabling Technologies for the Factories of the Future – Proceedings of the 48th CIRP Conference on Manufacturing Systems, pp. 69–74 (2016) https://doi.org/10.1016/j.procir.2015.12.006

Becker, T., Weimer, D.: Identification of autonomous structures in dynamic manufacturing networks using clustering approaches. In: The Annual Cambridge International Manufacturing Symposium. Disruptive Supply Network Models in Future Industrial Systems: Configuring for Resilience and Sustainability, Management Technology, London, pp. 1–11 (2014)

Becker, T., Meyer, M., Beber, M.E., Windt, K., Hütt, M.T.: A comparison of network characteristics in metabolic and manufacturing systems. Lect. Notes Ser. 141–150 (2013). http://dx.doi.org/10.1007/978-3-642-35966-8_11

Becker, T., Beber, M.E., Windt, K., Hütt, M.T., Helbing, D.: Flow control by periodic devices: a unifying language for the description of traffic, production, and metabolic systems. J. Stat. Mech: Theory Exp. **2011**(05), P05004 (2011). https://doi.org/10.1088/1742-5468/2011/05/p05004

Becker, T., Beber, M.E., Windt, K., Hütt, M.T.: The impact of network connectivity on performance in production logistic networks. CIRP J. Manuf. Sci. Technol. **5**(4), 309–318 (2012). https://doi.org/10.1016/j.cirpj.2012.09.004. Special issue from 44th CIRP Conference on Manufacturing Systems

Becker, T., Meyer, M., Windt, K.: A manufacturing systems network model for the evaluation of complex manufacturing systems. Int. J. Product. Perform. Manag. **63**(3), 324–340 (2014). https://doi.org/10.1108/ijppm-03-2013-0047

Blunck, H., Vican, V., Becker, T., Windt, K.: Improvement heuristics for manufacturing system design using complex network figures. In: Procedia CIRP, vol. 17, pp. 50–55 (2014). Variety Management in Manufacturing. https://doi.org/10.1016/j.procir.2014.01.063

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: structure and dynamics. Phys. Rep. **424**(4–5), 175–308 (2006). https://doi.org/10.1016/j.physrep.2005.10.009, http://www.sciencedirect.com/science/article/pii/S037015730500462X

Braha, D., Bar-Yam, Y.: From centrality to temporary fame: dynamic centrality in complex networks. Complexity **12**(2), 59–63 (2006). https://doi.org/10.1002/cplx.20156, https://onlinelibrary.wiley.com/doi/abs/10.1002/cplx.20156, https://onlinelibrary.wiley.com/doi/pdf/10.1002/cplx.20156

De Menezes, M.A., Barabási, A.L.: Fluctuations in network dynamics. Phys. Rev. Lett. **92**(2), 028701 (2004)

Funke, T., Becker, T.: Stochastic block models: a comparison of variants and inference methods. PLoS ONE **14**(4), e0215296 (2019). https://doi.org/10.1371/journal.pone.0215296.

Funke, T., Becker, T.: Complex networks of material flow in manufacturing and logistics: modeling, analysis, and prediction using stochastic block models. J. Manuf. Syst. **56**, 296–311 (2020). https://doi.org/10.1016/j.jmsy.2020.06.015, https://www.sciencedirect.com/science/article/pii/S0278612520301035

Gronau, N., Theuer, H.: Potenziale autonomer Technologien in Produktion und Logistik. Logistik: Gestern, heute, morgen – Festschrift für Richard Vahrenkamp zur Vollendung des **65**, 269–286 (2011)

Lämmer, S., Gehlsen, B., Helbing, D.: Scaling laws in the spatial structure of urban road networks. Phys. A Stat. Mech. Its Appl. **363**(1), 89–95 (2006)

Lancichinetti, A., Kivelä, M., Saramäki, J., Fortunato, S.: Characterizing the community structure of complex networks. PLoS ONE **5**, e11976 (2010). https://doi.org/10.1371/journal.pone.0011976

Lovász, L.: Random walks on graphs: a survey, combinatorics, paul erdos is eighty. Bolyai Soc. Math. Stud. **2**, 1–46 (1993)

Meepetchdee, Y., Shah, N.: Logistical network design with robustness and complexity considerations. Int. J. Phys. Distrib. Logist. Manag. **37**(3), 201–222 (2007). https://doi.org/10.1108/09600030710742425

Mourtzis, D., Doukas, M.: Decentralized manufacturing systems review: challenges and outlook. Robust Manufact. Control 355–369 (2013)

Omar, Y.M., Minoufekr, M., Plapper, P. Lessons from social network analysis to industry 4.0. Manuf. Lett. **15**, 97–100 (2018)

Rosvall, M., Bergstrom, C.T.: Maps of random walks on complex networks reveal community structure. Proc. Natl. Acad. Sci. **105**(4), 1118–1123 (2008). https://doi.org/10.1073/pnas.0706851105, https://www.pnas.org/content/105/4/1118, https://www.pnas.org/content/105/4/1118.full.pdf

Schaeffer, S.E.: Graph clustering. Comput. Sci. Rev. **1**(1), 27–64 (2007). https://doi.org/10.1016/j.cosrev.2007.05.001, https://www.sciencedirect.com/science/article/pii/S1574013707000020

Vrabič, R., Husejnagić, D., Butala, P.: Discovering autonomous structures within complex networks of work systems. CIRP Annal. **61**(1), 423–426 (2012). https://doi.org/10.1016/j.cirp.2012.03.083, https://www.sciencedirect.com/science/article/pii/S0007850612000856

Vrabič, R., Škulj, G., Butala, P.: Anomaly detection in shop floor material flow: a network theory approach. CIRP Ann. Manuf. Technol. **62**(1), 487–490 (2013). https://doi.org/10.1016/j.cirp.2013.03.131, http://www.sciencedirect.com/science/article/pii/S0007850613001327

Wagner, D., Becker, T.: Characterization of cluster structures in material flow networks: a network approach. Procedia Technol. **26**, 277–283 (2016). 3rd International Conference on System-Integrated Intelligence: New Challenges for Product and Production Engineering. https://doi.org/10.1016/j.protcy.2016.08.037, https://www.sciencedirect.com/science/article/pii/S221201731630384X

Wagner, D., Becker, T.: Evaluation of the applicability of random walks for generation of material flow network models. Proc. CIRP **63**, 488–492 (2017). Manufacturing Systems 4.0 Proceedings of the 50th CIRP Conference on Manufacturing Systems. https://doi.org/10.1016/j.procir.2017.03.164, https://www.sciencedirect.com/science/article/pii/S2212827117303463

Wagner, D., Becker, T.: Automated generation of material flow network models with cluster structures using random walks. Proc. CIRP **72**, 569–573 (2018). 51st CIRP Conference on Manufacturing Systems. https://doi.org/10.1016/j.procir.2018.03.249, https://www.sciencedirect.com/science/article/pii/S2212827118304207

Williams, R.J., Berlow, E.L., Dunne, J.A., Barabási, A.L., Martinez, N.D.: Two degrees of separation in complex food webs. Proc. Natl. Acad. Sci. **99**(20), 12913–12916 (2002)

Recent Developments in Mathematical Traffic Models

Daniel Schmand

Abstract Predictions such as forecasts of congestion effects in transportation networks can be based on complex simulations that include many aspects of actual transportation systems. On the other hand, rigorous mathematical traffic models give rise to theoretical analyses, very general statements, and various traffic optimization opportunities. There has been a huge development in the last years to make mathematical traffic models more realistic. This chapter provides an overview of the mathematical traffic models developed recently and some state-of-the-art results.

1 Introduction

One of the major challenges in logistics is the steadily growing traffic activity. The European Commission expects that passenger transport activities across Europe will increase by 42% by 2050 and freight transport by 60% (Mobility and Transport 2019). This puts pressure on the capacity of the available transport network since congestion leads to high associated costs, such as lost time, increased vehicle operating costs, and environmental aspects. The Department of Transport, Tourism and Sport's Economic and Financial Unit of the Republic of Ireland estimated the cost of time lost due to aggravated congestion in the Greater Dublin Area to be €358 million in 2012. They forecast a cost of €2.08 billion per year in 2033 (Department of Transport, Tourism and Sport 2017). Ensuring that congestion does not reach an unacceptable level is of significant high importance for all future logistics processes.

Climate change is one of the main challenges of our times. In 2019, global CO_2 emissions increased to 38 gigatons due to a report by the Netherlands Environmental Assessment Agency (Olivier and Peters 2020). The transport sector is still one of the major contributors to greenhouse gas emissions, and congestion increases the emissions of greenhouse gases of vehicles. Additionally, congestion imposes a negative impact on local noise and water quality (Department of Transport, Tourism

D. Schmand (✉)
Center for Industrial Mathematics, University of Bremen, Bremen, Germany
e-mail: schmand@uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_4

and Sport 2017). As a consequence, reducing congestion does not only reduce travel times and reduce the cost of time for logistics operators, but it also has positive effects on greenhouse gas emissions and quality of life.

One of the first steps to reduce congestion and increase the efficiency of networks is to understand the behavior of traffic. Typically, researchers evaluate traffic by creating a theoretical model that aims to capture as many aspects of real-life traffic as possible. We can differentiate the existing traffic models into macroscopic, mesoscopic, and microscopic models. Macroscopic models focus on the joint behavior of vehicle flows, and thus it is usually not possible to track a single vehicle. Examples of macroscopic models are function-based network load models. Microscopic models usually track single vehicles in a highly temporal solution. They capture microscopic aspects like overtaking, traffic lanes, human driver behavior, and others. Mesoscopic models can be seen as hybrids. For a nice comparison of the models, we refer the reader to the survey of Wang et al. (2018). Usually, macroscopic models are easier to understand and easier to use in software but typically lack precision and many local aspects of real-life traffic. The traffic simulation software MATSim is one of the state-of-the-art implementations of a microscopic to mesoscopic model; we refer to *the MATSim book* for a great introduction (Horni et al. 2016). Individual vehicles can be tracked and followed, and crossings are implemented from a microscopic point of view, but the actual behavior of vehicles on a road segment is not modeled precisely. The current implementation can be roughly described as follows. The traffic network is modeled by a graph which consists of vertices and edges. A vertex corresponds to a crossing. The directed edges connect vertices and model the road segments between crossings. Vehicles that enter some edge at some point in time are instantaneously moved to the end of the edge and have to wait there for at least the externally given free-flow travel time of that edge (Horni et al. 2016). At the head of the edge, a capacity constraint is checked, and the vehicle is only allowed to leave the edge if the capacity is not exceeded. The software drops some precision in modeling the microscopic behavior on an edge with the aim to reduce computation time and thus increase the size of the models that can be analyzed.

Microscopic to mesoscopic simulations are a great tool to get insights on traffic behavior, but usually one can only observe properties of the model by simulations, and it is very hard to prove properties of the model or the solution. To illustrate this, consider MATSim's co-evolutionary algorithm. Usually, MATSim simulates traffic for 1 day. The simulation is done in multiple iterations. Agents in the simulation have a list of *plans* (although plans capture more, it is easiest to think of different plans being different possible routes). In each iteration, each agent picks a randomly chosen plan from the list with a probability that depends on the current *score* of this plan. After each agent has chosen a plan, the actual day is simulated and the agents evaluate their plan and update the score of the chosen plan with respect to the observed traffic. The updated score is used in the next iteration. Additionally, in each iteration, some agents (usually $\sim$10%) are allowed to create new plans. The process is repeated until the average population score does not change more than some threshold. For more details, we refer the reader to *the MATSim book* (Horni et al. 2016). A main drawback of this model is that rigorously proving properties both for the process and for this final state is a very challenging task. Sering et al. (2021) recently showed first provable results on the network loading model in MATSim, but even the question if there is a best response dynamic that converges to an approximate equilibrium is still open. This implies that rigorously proving properties of the resulting states is a very challenging task, and we cannot use theoretical insights to improve the results, e.g., by trying to enforce a provable optimal traffic distribution. Generally, microscopic to mesoscopic models can currently only give an estimation of what might happen to, e.g., a change of the input of other questions of interest, and the obtained results cannot be justified analytically.

Macroscopic models on the other hand can be analyzed mathematically but lack many important properties of traffic flows. Naturally, researchers aim to set up mathematical models that cover as many properties of traffic flows as possible and are still analyzable. There was a huge breakthrough in this direction in the last years, and the aim of this chapter is to cover this development. One of the most important differences in mathematical traffic models is the difference between discrete and continuous flows. Discrete flows are a model with indivisible particles of a certain size. The motivation and connection to traffic are immediately clear, and some of these models also work with particles of different sizes. On the other hand, continuous models treat traffic as divisible in arbitrarily small pieces. Another very important property of mathematical traffic models is the existence of strategic users. Typically, traffic flow cannot be controlled by a central authority, and the users strategically decide which route to take. This behavior can be seen as the mathematical analog to the agents that individually (and randomly) choose their best plans in MATSim. Note that in this work, we are interested in *dynamic* flows, which in contrast to their static counterparts implement a time component. Thus, the network can have different states at different times. There is a very rich literature on static strategic routing models (i.e., without a time component), which we will not cover in this chapter. For the readers interested in static models, we refer to the terms *Wardrop games*, *network congestion games*, and *(non-) atomic selfish routing*.

The rest of the chapter is divided into two parts. Section 2 covers the continuous model and Sect. 3 the discrete variants.

2 Continuous Flows: Nash Flows Over Time

Vickrey (1969) was the first to describe a basic model for continuous flows over time. The model captures the time-dependent behavior of arbitrarily splittable flow in networks with edges with a limited throughput capacity. This model is typically called *dynamic flows with deterministic queuing*. Koch and Skutella extended this model and added strategic user's behavior (Koch and Skutella 2011). Their model is called *Nash flows over time*. In recent years, Cominetti, Correa, Cristi, Olver, Oosterwijk, and many others proved many properties of flows in this model, see, e.g., Cominetti et al. (2011, 2015, 2017); Correa et al. (2019). Sering and Vargas Koch achieved a huge breakthrough by adding spillback to this setting. In their extension, traffic users need some space for queuing, and thus edges can become full (Sering and Vargas Koch 2019). In this sense, highly congested parts of the network now can have an impact on previous edges as traffic users might spill back to other parts of the network while queuing. Additional properties of this model have been shown by Israel and Sering (2020). Ziemke et al. recently presented experiments that indicate a strong connection of the limit of the MATSim flow model for decreasing vehicle and time step size and Nash flows over time with spillback (Ziemke et al. 2021). Their article provides a strong justification and motivation for the studies of mathematical traffic models.

We will proceed by describing the ideas of the basic model by Vickrey. The model definition is mainly based on the works of Koch and Skutella (2011), Cominetti et al. (2011), and Sering (2020). In the subsequent subsections, we present the extensions and some results established in the literature.

2.1 Continuous Flows Over Time

The most important ingredient to the flow over time model is the edge dynamics. Road networks are modeled by a graph $G = (V, E)$ with vertices V and directed edges E. Each edge $e \in E$ is equipped with a (free-flow) travel time $\tau_e \geq 0$ and outflow capacity $\nu_e > 0$. Usually, it is assumed that there are no cycles of length 0 in G. The inflow rate into an edge can be arbitrarily large and is not restricted. However, if the inflow rate exceeds the capacity at some point in time θ, the outflow rate at $\theta + \tau_e$ is restricted to ν_e due to the outflow capacity of the edge. All additional flow is stored in a *point queue*, or sometimes called *horizontal queue*. In this sense, each edge can store an arbitrary amount of flow in a queue and can never get full. If there is a queue at some edge, first the flow waiting in the queue is allowed to leave the edge, i.e., in this sense, there is a FIFO property in each edge. Put differently, overtaking is not possible. The total mass that is stored in the queue of edge e at time θ is denoted by $z_e(\theta)$. Figure 1 is an illustration of the basic notions of the edge dynamics.

2.1.1 Connectors of the Edges: Vertices

Edges are connected to each other by vertices and thus vertices model crossings. They only connect edges and flow leaving an edge e_1 is just forwarded to the subsequent edges. This will change drastically in the later parts of the section.

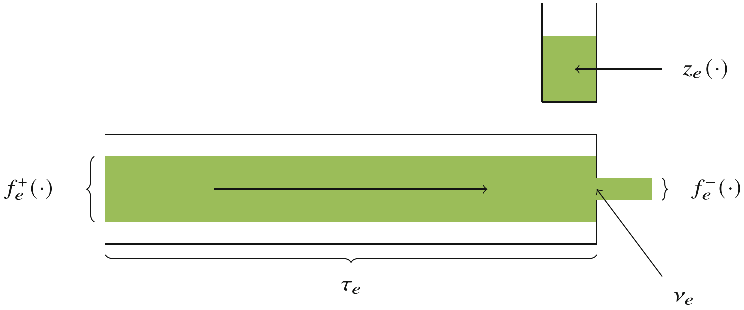


Fig. 1 An illustration of the basic notions of the edge dynamics

2.2 Mathematical Consistency

In order to set up the model precisely, we need to formalize all properties. Formally, a flow over time is defined by a family of functions $f = (f_e^+, f_e^-)_{e \in E}$, where f_e^+ and f_e^- are integrable nonnegative functions and specify the flow rate entering or leaving edge e at a given time, respectively. We can calculate

$$F_e^+(\theta) = \int_0^\theta f_e^+(\xi)\, d\xi \qquad \text{and} \qquad F_e^-(\theta) = \int_0^\theta f_e^-(\xi)\, d\xi \; .$$

Note that in this chapter, we will enforce all relevant conditions for all times $\theta > 0$. In the original literature, they are only enforced for almost all times θ, i.e., typically they are allowed to be not fulfilled for a set of times that forms a null set. In this chapter, we aim for giving the main ideas and decided to lack this mathematical detail as it simplifies arguments and notations. We refer the reader to the original works for more details.

We say that flow conservation on a vertex $v \in V$ is fulfilled if

$$\sum_{e \in \delta_+(v)} f_e^-(\theta) = \sum_{e \in \delta_-(v)} f_e^+(\theta) \; ,$$

for all θ. Here, $\delta_+(v)$ and $\delta_-(v)$ denote the incoming and outgoing edges from vertex v, respectively. Thus, flow conservation at v is fulfilled if for all times θ the total flow rate entering vertex v is equal to the total flow rate leaving v. We do not enforce flow conservation on *all* vertices. In the basic model, there is a single designated vertex $s \in V$, called source, for which

$$\sum_{e \in \delta_+(s)} f_e^-(\theta) + r = \sum_{e \in \delta_-(s)} f_e^+(\theta) \; ,$$

for some externally given constant inflow rate r into s. Additionally, there is one sink $t \in V$ for which we only enforce

$$\sum_{e \in \delta_+(v)} f_e^-(\theta) \geq \sum_{e \in \delta_-(v)} f_e^+(\theta) \ .$$

These constraints ensure that flow enters the system with a constant flow rate r at $s \in V$, is maintained in the system at intermediate vertices $v \in V \setminus \{s, t\}$, and is allowed to leave the system at t.

The amount of flow that is stored in the queue at the head of edge $e \in E$ at time $\theta + \tau_e$ is the amount of flow that has already entered edge e by θ and not left e by time $\theta + \tau_e$. It can be calculated by

$$z_e(\theta + \tau_e) = F_e^+(\theta) - F_e^-(\theta + \tau_e) \ .$$

Now, we have all ingredients to define a feasible flow over time. We say a flow $f = (f_e^+, f_e^-)_{e \in E}$ is a feasible flow over time with source s, inflow rate r, and sink t if it fulfills flow conservation on all vertices $v \in V \setminus \{s, t\}$, the above conditions for s and t, and the following property. The outgoing flow rates f_e^- have to fulfill

$$f_e^-(\theta + \tau_e) = \begin{cases} \nu_e & \text{if} z_e(\theta + \tau_e) > 0 \ , \\ \min\{\nu_e, f_e^+(\theta)\} & \text{otherwise} \ , \end{cases}$$

where ν_e denotes the outflow capacity of edge e. This condition also ensures that flow is not created inside edges, i.e.,

$$F_e^-(\theta + \tau_e) \leq F_e^+(\theta) \ ,$$

for all times $\theta \geq 0$ and all edges $e \in E$. Note that for given flow rates f_e^+, the flow rates f_e^- are uniquely defined and can be calculated according to the equation above.

The flow in this model can be any integrable nonnegative function, and thus flow can be divided arbitrarily and this is a continuous flow model. On the other hand, we aim for traffic models and are particularly interested in the behavior of single vehicles. In this model, one can interpret the flow as infinitely many infinitesimal small particles. Given some feasible flow over time, we are typically interested in the travel time of some particle from s to t. In order to calculate that, we need to track the chosen path and the observed waiting time in queues for this particle. We follow standard notation in the literature and define the waiting times of a particle entering edge e at time θ as

$$q_e(\theta) = \frac{z_e(\theta + \tau_e)}{\nu_e} \ .$$

Analogously, one can calculate the exit time of a particle entering e at θ by

$$T_e(\theta) = \theta + \tau_e + q_e(\theta) \ .$$

2.3 *The Nash Condition*

The model described so far is a well-defined mathematical model for traffic. Given some road networks with edge capacities and free-flow travel time, we can build a mathematical twin of a traffic network. This model respects over-time behavior and assumes that traffic flow is arbitrarily splittable, i.e., we expect the model to be most accurate with an increasing number of users that individually have a neglectable contribution to congestion. Given some inflow into a certain vertex and specified paths for the traffic users, we can set up a feasible flow over time respecting the setup and the flow conservation constraint. Generally, in traffic optimization, we are mostly interested in a traffic model that can actually predict traffic, while we are only given the network and a certain demand, represented by the inflow (and desired outflow). In order to do so, we need to define a system that describes how the paths of the traffic particles in the network are chosen. In the following, we will follow the road of Koch and Skutella (2011) and extend the model with strategic users.

We assume that flow particles always try to individually minimize travel time to the destination and choose a currently shortest path with respect to the other flow particles. Formally, given some flow f, for a particle entering the source at time ϕ, we define $l_v(\phi)$ to be the earliest possible arrival time at vertex v. With $T_e(\theta)$ being the exit time on e of a particle entering e at θ, this imposes the characterization

$$l_v(\phi) = \min_{(u,v)\in\delta_+(v)} T_{(u,v)}(l_u(\phi)) \ ,$$

for all $v \in V \setminus \{s\}$ and $l_s(\phi) = \phi$. We now set up the *current shortest path network* as follows. We say an edge $(u, v) \in E$ is active for entering time ϕ and flow f if $l_v(\phi) = T_{(u,v)}(l_u(\phi))$. These are exactly the edges that can possibly be used by a particle entering at time ϕ on a shortest path. The current shortest path network $G'_f(\phi) = (V, E'_f(\phi))$ contains exactly all active edges for ϕ and f, i.e.,

$$E'_f(\phi) = \left\{(u, v) \in E \mid l_v(\phi) = T_{(u,v)}(l_u(\phi))\right\} \ .$$

Definition 1 (Based on Koch and Skutella (2011)) A flow over time f is called a dynamic equilibrium or a Nash flow over time if

$$f^+_{(u,v)}(l_u(\phi)) > 0 \Rightarrow (u, v) \in E'_f(\phi) \ ,$$

for all edges $(u, v) \in E$ and departure times ϕ.

Equivalently, all particles only use edges that are part of their current shortest path network. Cominetti, Correa, and Larré established the following property of Nash flows over time.

Theorem 1 (Based on Cominetti et al. (2011)) *A flow over time f is a Nash flow if and only if*

$$F_e^+(l_u(\phi)) = F_e^-(l_v(\phi))$$

for all edges $e = (u, v) \in E$ and departure times ϕ.

Note that in the original works, both Definition 1 and Theorem 1 are only valid for almost all times θ, so they allow for null sets of times for which the conditions are not true. Informally, the theorem assures that the Nash flow condition is equivalent to the fact that exactly all flow that has entered edge (u, v) by time $l_u(\phi)$ has left the edge by time $l_v(\phi)$. This ensures that, for a particle entering the network at ϕ, we can define $x_e(\phi) = F_e^+(l_u(\phi))$ for all edges. It can be shown that $(x_e(\phi))_{e\in E}$ is a static flow with flow value $r\phi$, where r denotes the constant inflow rate, see, e.g., Cominetti et al. (2011). Due to the definition of x, the flow value $x_e(\phi)$ measures exactly the amount of flow that has entered edge e before particles starting at ϕ can reach e. The derivative $x_e'(\phi)$ (if it exists) is also a static flow with flow value r and given by

$$x_e'(\phi) = f_e^+(l_u(\phi)) \cdot l_u'(\phi) \ .$$

The term $l_u'(\phi)$ can be interpreted as the factor incorporating the *shift of time* present at u, i.e. the behavior of $l_u(\phi)$ in contrast to ϕ. In the subsequent section, we describe the main ideas for constructing a Nash flow over time by making use of this characterization. Throughout the construction, we will make sure that the derivatives l_u' do exist.

2.4 Constructing Nash Flows Over Time

Koch and Skutella (2011) were the first to describe the existence of Nash flows over time. Cominetti, Correa, and Larré improved and extended their results (Cominetti et al. 2011). Here, we will only give an informal description and the main ideas of the constructive procedure described by Cominetti et al. (2011). We refer the reader to the original work or to the nice description of the procedure in Sering's PhD thesis (Sering 2020) for more mathematical details.

The main idea builds on the fact that there is always a Nash flow over time that can iteratively be constructed in phases. In order to do so, we split the set of edges into three disjoint sets. Let $E_f'(\phi)$ denote the set of active edges, $E_f^*(\phi) \subseteq E_f'(\phi)$ be the set of *resetting edges*, i.e., those edges $(u, v) \in E_f'(\phi)$ for which there is a non-empty queue at time $l_u(\phi) + \tau_e$, and the inactive edges $E \setminus E'_f(\phi)$. It turns out that these three sets of edges define the phases of the Nash flow and that we can define a static flow x' corresponding to the static flow x' defined earlier, which does not change during a phase. For a proof of this statement, see Cominetti et al. (2011). In the following, we will fix some phase and give a description of constraints for the static flow x' in this phase.

We define the phases with respect to starting times of particles. Let some phase cover the particles leaving s in the interval $[\phi_1, \phi_2)$ and x' be the corresponding static flow for this phase. The values for ϕ_1 and ϕ_2 will be chosen such that the sets of active and resetting edges do not change within a phase. As the particles leaving s in the interval $[\phi_1, \phi_2)$ can only arrive at some edge (u, v) during time $[l_u(\phi_1), l_u(\phi_2))$, this is the interesting time interval for the edge for this phase. However, if the inflow $f^+_{(u,v)}$ into some edge (u, v) is constant during the phase, the outflow corresponding to the phase (and thus shifted by the time needed to traverse the edge) is given as follows. The outflow is equal to ν_e if the edge e has a queue of positive size and f^+_e otherwise. This ensures constant outflow during a phase, and thus we establish constant inflow and outflow for all edges. Given this, the change of the queues is also constant on all edges, and the idea is to construct a static flow x' such that all flow particles only use active edges and to distribute the flow particles such that the change of the total waiting time is the same for all used paths and subpaths. Then, $l'_u(\theta)$ exists for all $\theta \in (\rho_1, \rho_2)$, the right derivative exists at ρ_1, and all are equal.

Due to the desired behavior of x' and l', we have to make sure that x' is chosen such that $l'_s = 1$ and

$$l'_v = \min_{e=(u,v)\in E'} \rho_e(l'_u, x'_e) ,$$

where

$$\rho_e(l'_u, x'_e) = \begin{cases} \frac{x'_e}{\nu_e} & \text{for } e \in E^* , \\ \max\left\{l'_u, \frac{x'_e}{\nu_e}\right\} & \text{otherwise} . \end{cases}$$

After the calculation of x'_e and l'_u, we can define

$$f^+_e(\theta) = \frac{x'_e}{l'_u}$$

for all $\theta \in [l_u(\rho_1), l_u(\rho_2))$. This ensures that the static flow is an equilibrium not only for particles starting at time ϕ_1 but also for all later particles in the phase, since the travel times change on all used paths and subpaths by the same amount. Koch and Skutella have already introduced these static flows, and Cominetti, Correa, and

Larré have extended their results. For more mathematical details, we refer to the original works (Cominetti et al. 2011; Koch and Skutella 2011).

We construct the equilibrium phase by phase starting at time 0. Clearly, a phase ends if a new edge enters the current shortest path network or if some queue depletes. Note that this procedure by Cominetti et al. (2011) constructively shows existence of Nash flows over time. We conclude the following theorem.

Theorem 2 (Cominetti et al. (2011)) *There is always a Nash flow over time.*

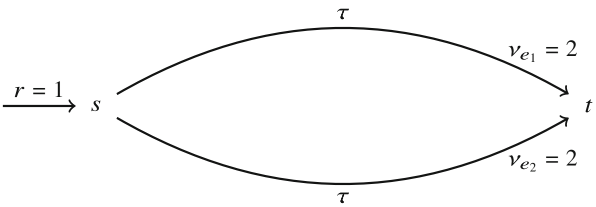
However, the procedure does not give rise to a polynomial time algorithm. Moreover, one of the main open problems is to answer the question if the number of phases is bounded.

2.5 Analysis of Equilibria

Equilibria in general and also those that are constructed according to the method above are not unique. As an easy example, consider a graph with two parallel edges with equal travel time τ and capacity $\nu_{e_1} = \nu_{e_2} = 2$ and an inflow rate of $r = 1$, cf. Fig. 2. As it turns out, there are no queues and all static flows are feasible for the first (and only) phase. The values $l_t(\phi)$ are equal to $\phi + \tau$ and coincide for all choices. The particles leaving at time ϕ still arrive at the same time in all possible choices. More generally, Cominetti, Correa, and Larré have shown that the arrival time functions $l_v(\phi)$ are equivalent as long as we restrict ourselves to Nash flows over time f with right-continuous inflow and outflow rates (Cominetti et al. 2011). It is not proven formally that this is true for all equilibria, but it is conjectured to be true (Sering 2020).

In order to evaluate the existing equilibria or Nash flows in a network, the notions *price of anarchy* and *price of stability* have been established. Usually, one defines an objective function from the central's authority's point of view, and the objective function value of an equilibrium solution is compared to the solution value of an optimal solution with respect to this objective function. If all equilibria in the same network have the same objective function value, the prices of anarchy and stability coincide. As this is currently conjectured for Nash flows over time, we will focus on results on the price of anarchy. The price of anarchy is defined as the worst possible objective function value of a Nash flow divided by the optimal solution value in this

Fig. 2 An example showing that Nash flows over time are not unique



instance. For flows over time, two different objective functions have been analyzed in the literature, see Bhaskar et al. (2015).

For the first measure, we assume that there is a certain amount of flow volume A that has to be sent from s to t. The *makespan price of anarchy* measures the arrival time of the latest particle. This can be useful in a situation, where we are interested in sending the total flow volume as fast as possible, e.g., in some evacuation scenarios.

The second measure assumes that we are given a time horizon and are allowed to send as much flow as we can from s to t before the deadline. This *throughput price of anarchy* measures the total amount of flow that arrives at t before the deadline. This can be useful for many logistics processes, where we are interested in sending as much flow as possible in a given time.

For Nash flows over time, it is shown that the throughput price of anarchy is unbounded in the network size.

Theorem 3 (Koch and Skutella (2011)) *The throughput price of anarchy lies in* $\Omega(|E|)$.

Correa, Cristi, and Oosterwijk showed that the makespan price of anarchy is upper bounded by $\frac{e}{e-1}$ if the following conjecture is true. Consider the same graph with two different inflow rates $r_1 > r_2$. Let f_1 be a Nash flow over time for inflow r_1 and f_2 be a Nash flow over time for inflow r_2. Then, $l_v^{f_1}(\frac{A}{r_1}) \leq l_v^{f_2}(\frac{A}{r_2})$, i.e., a faster arrival of the same particles does not induce a later arrival time for the latest of these particles (Correa et al. 2019).

Theorem 4 (Correa et al. (2019)) *Under the assumption that a higher inflow does not lead to a later arrival time at the sink of the last particle of a constant size flow mass, the makespan price of anarchy is at most* $\frac{e}{e-1}$.

From a traffic planner's point of view, the price of anarchy is a measure of the inefficiency of the network due to selfishness and the individual objective functions of the users. In this sense, it tells the network designer how much efficiency is lost in the worst case equilibrium, because there is no central authority that controls all vehicles.

One of the main open problems in the area is to prove or disprove the conjecture that the assumption in Theorem 4 is always true.

2.6 Spillback

Sering and Vargas Koch made a huge breakthrough and extended the model by storage capacities and inflow capacities on the edges. The storage capacities represent the number of particles that can wait on an edge and induce an important property of real-life traffic into the model. As soon as the storage capacity is reached, the inflow cannot be larger than the outflow of the edge. This imposes spillback behavior as waiting particles are no longer local but might have a global effect. In order to extend the procedure to calculate a Nash flow over time in the spillback model, they added two additional events that introduce new phases. The static flow can only remain constant as long as the set of full edges does not change, the outflow on an edge does not change, and the sets of active and resetting edges remain constant. With this change to the procedure, they showed the following result.

Theorem 5 (Sering and Vargas Koch (2019)) *There is always a Nash flow over time with spillback.*

If some edge becomes full, it still accepts flow equal to the current outflow of this edge. Sering and Vargas Koch assumed a *fair allocation condition*, that is, flow entering an edge merges proportional to the outflow capacity of the previous edges. The authors only use this assumption for establishing the spillback model, but it can also be used to implement traffic lights in the mathematical model. In practice, most traffic lights run in the so-called *cycles*. Each cycle has a certain length and each lane gets at once a green light in a cycle. The *split time* determines for how long the lights should stay green. Main roads can get preference by a longer split time in the corresponding phase. From a global point of view, traffic lights can be seen as splitting up the inflow capacities of a certain edge to the incoming edges. However, by limiting the inflow capacity of a certain edge and setting the outflow capacities according to the traffic lights, the model by Sering and Vargas Koch does also provide features to mimic traffic lights.

2.7 Further Notes and Remarks

The procedure that we described here to calculate Nash flows over time does not give rise to a polynomial time algorithm to calculate the static flows of the phases. It was shown by Sering (2020) that one can set up a mixed-integer program (MIP) to calculate the static flows and the end of the phases in the basic model. The MIP formulation for the spillback model can be found in Sering and Vargas Koch (2019). Sering and Zimmer published a tool that is based on the MIP formulations and can be used to calculate Nash flows over time (Sering and Zimmer 2020), even in the spillback model.

Cominetti et al. (2017) analyzed the long-term behavior of Nash flows over time in the basic model. They were interested in the studies of the last phase of a Nash flow over time. They showed that there is a steady state, i.e., an infinitely long last phase in which the travel time remains constant for all particles if and only if the inflow rate does not exceed the minimum cut capacity of the network. This might be particularly interesting from a traffic designer's point of view.

So far, we have only described models with networks with a single source and a single sink vertex. Cominetti et al. (2015) showed that the basic model can also be extended to a multi-commodity network, i.e., to a network with multiple source–sink pairs while still possessing a Nash flow over time.

Mathematical traffic models are already capable of modeling capacitated roads, congestion, individual particles' objectives, spillback effects, and traffic lights. In this sense, the models are quite accurate, but mathematical results are mainly established for single source–single sink networks, which are quite limited for applications. There are results on the price of anarchy and the efficiency of equilibria. A main aspect for future work could be the question if it is possible to change the network by, e.g., artificially limiting the capacity of certain roads or by introducing tolls to improve the equilibrium solution. Researchers have already shown similar results in the static (i.e., non-over-time) setting, see, e.g., Beckmann et al. (1956), Dafermos and Sparrow (1969), and it would be an interesting and promising direction to establish similar results for dynamic traffic models. Results of this kind could potentially be applied to real-life traffic and reduce congestion in transportation networks, i.e., improve on logistics services.

3 Discrete Flows: Competitive Packet Routing

The second research direction in mathematical traffic models is models for discrete flows over time. In contrast to the continuous model, individual indivisible users are present. Dependent on the exact definition of the model, users may be weighted. Werth, Holzhauser, and Krumke were the first to introduce a discrete variant of Koch and Skutella's Nash flows over time (Werth et al. 2014). The model was later called competitive packet routing (Harks et al. 2018) as it can also be seen as the competitive variant of the well-studied packet routing model.

The main idea is to use integral travel times. Each edge has some certain outflow capacity and only users with an aggregated weight smaller or equal to the edge capacity can leave the edge at each integral point in time. Those who are not allowed to leave the edge have to queue and wait.

3.1 The Mathematical Model

An instance of the problem is given by a graph $G = (V, E)$, integral travel times τ_e, capacities ν_e for all $e \in E$, and n users, each with a start vertex $s_i \in V$, a target vertex $t_i \in V$, and a weight w_i for all $i \in \{1, \dots, n\}$. Typically, an instance is called unweighted if $w_i = 1$ for all $i \in \{1, \dots, n\}$. A strategy profile $P = (P_1, \dots, P_n)$ specifies a path for each user, where for P_i we only allow a simple path from s_i to t_i such that $\nu_e \geq w_i$ for all edges $e \in P_i$.

The edge's behavior and the vertices are set up using discrete time steps. A user that arrives at some vertex $u \in V$ at an integral time θ can immediately enter the next edge $e = (u, v)$ on the specified path. Then, we assume that the user needs τ_e time units to travel along the edge and reaches the queue Q_e at the head of the edge at time $\theta + \tau_e$. After adding users to the queues at time $\theta + \tau_e$, the first users from Q_e with maximum aggregated weight not exceeding the capacity ν_e can leave edge e and enter vertex v. However, in order to define the model properly, one needs to specify rules that describe which users are allowed to leave the edge if not all can leave the edge due to capacity issues. Werth, Holzhauser, and Krumke assume that the users are processed in the order of arrival at the edge (Werth et al. 2014). However, if multiple users enter the edge at the same time, there is a tie-breaking according to some priority list. For tie-breaking, Werth, Holzhauser, and Krumke use either a global order of users valid in the whole network or a priority order depending on the edge from which the users have entered the questionable edge.

They consider two different objective functions, namely a sum and a bottleneck objective. For the bottleneck objective, users aim for minimizing the maximum time they spend on an edge and the social cost is given by the maximum value of all users. As this objective does not seem closely related to traffic, we focus on the sum objective. In this variant, the users aim to minimize the sum of travel times on their path and the social objective is given by the makespan, i.e., by the largest arrival time of all users. Werth, Holzhauser, and Krumke show that Nash equilibria exist in the case that all users share the same source and smaller users have a higher priority than larger users. If this is not the case, they showed that Nash equilibria might not exist.

3.2 *Extensions of the Model*

Scarsini et al. (2018) analyzed a model equivalent to the one by Werth, Holzhauser, and Krumke for users that arrive at the source node sequentially. Their aspect was on how the system evolves to a steady state. Harks et al. considered almost the same model but with other objective functions (Harks et al. 2018). In their work, the social objective is the sum of the user's arrival times (in contrast to the maximum in the work by Werth, Holzhauser, and Krumke). Users are unweighted and at a queue of some edge users are either prioritized with respect to a given global order of users or with respect to a local order of users that is only valid at this edge. This is in contrast to the FIFO setup in the original model, and it is the only model covered here that allows users to overtake other users on the same path. Harks et al. show that Nash equilibria exist in networks with global priorities (Harks et al. 2018). The idea is that if some user i has higher priority than some other user j in all parts of the network, i can never be delayed by j. Thus, greedily assigning the users to their best responses in the priority order constructs a Nash equilibrium. Additionally, they gave various results on the prices of stability and anarchy for both local and global priorities that depend either on the number of users and edge capacities or on the length of a longest path. For more details, we refer the reader to their work (Harks et al. 2018).

Scheffler et al. (2018) modified the model by Harks et al. (2018) and studied their setting under edge priorities. These edge priorities are motivated by traffic rules and are a way more natural setup for traffic. They also provided the existence of equilibria for single-commodity instances and showed bounds on the prices of stability and anarchy. They provided examples where Nash equilibria contain cycles, which might also happen in real-world traffic. Cao et al. extended their model by adding the concept of *dynamic route choices*, i.e., users only choose their path while traveling (Cao et al. 2017). At each crossing, they decide on the next edge to take, instead of having to fix the whole path before starting to travel. This adaptive route choice is motivated by the replanning of routes while using GPS systems. Ismaili introduced the same behavior back to the original model of Werth, Holzhauser, and Krumke (Ismaili 2017).

Tauer et al. presented results on a train routing model which is based on the discrete store-and-forward packing routing model (Tauer et al. 2020). In their model, the users occupy space and deadlock and spillback effects are in place. Unfortunately, they do not consider strategic users, as trains can usually be routed by a central authority. Peis et al. considered a model for discrete flows over time, where some users might cooperate (Peis et al. 2018).

Hoefer et al. (2011) considered a slightly different variant than all the models mentioned above. They studied a variant with capacities on the number of users *that are on an edge* at the same point in time instead of having a constraint on the inflow or outflow of an edge. Additionally, their model works in a continuous-time setting.

3.3 Summary

Up to date, competitive packet routing models contain capacitated roads, congestion, individual particles' objectives, right of way rules, and model discrete indivisible users. Unfortunately, Nash equilibria are generally not unique, and the price of anarchy is quite large (in most cases linearly in the number of users). Additionally, spillback effects are usually not present, which is a huge drawback for traffic modeling. Although the variants with discrete users seem to be the more accurate model in the first place, it widely opens to understand spillback effects in the discrete model. The role and effects of discretization of time steps are often unclear. Sering et al. (2021) recently established first results in this direction and showed an interesting connection between continuous and discrete models. They proved that for decreasing packet sizes and time steps, the packet routing model converges to a flow over time model. In the discrete models, there are instances where different equilibria lead to huge differences in the travel times and it would be interesting to see in future work how much this is in line with real-world traffic. Similar to the continuous version, almost all discrete variants do not give the opportunity for users to overtake other users. An additional main aspect for future work could be the question of modifying the network in order to improve the quality of equilibria. This could be a first step in order to apply the gained knowledge to traffic networks and thus to logistics.

References

Beckmann, M., McGuire, C.B., Winsten, C.B.: Studies in the economics of transportation. Technical Report (1956)

Bhaskar, U., Fleischer, L., Anshelevich, E.: A stackelberg strategy for routing flow over time. Games Eco. Behav. **92**, 232–247 (2015)

Cao, Z., Chen, B., Chen, X., Wang, C.: A network game of dynamic traffic. In: Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, pp. 695–696. Association for Computing Machinery, New York (2017). https://doi.org/10.1145/3033274.3085101

Cominetti, R., Correa, J.R., Larré, O.: Existence and uniqueness of equilibria for flows over time. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) Proceedings of the Automata, Languages and Programming – 38th International Colloquium, ICALP 2011, Zurich, Switzerland, July 4–8. Part II, Lecture Notes in Computer Science, vol. 6756, pp. 552–563. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-22012-8_44

Cominetti, R., Correa, J.R., Larré, O.: Dynamic equilibria in fluid queueing networks. Oper. Res. **63**(1), 21–34 (2015). https://doi.org/10.1287/opre.2015.1348

Cominetti, R., Correa, J.R., Olver, N.: Long term behavior of dynamic equilibria in fluid queuing networks. In: Eisenbrand, F., Könemann, J. (eds.) Proceedings of the Integer Programming and Combinatorial Optimization – 19th International Conference, IPCO 2017, Waterloo, ON, June 26–28. Lecture Notes in Computer Science, vol. 10328, pp. 161–172. Springer, Berlin (2017). https://doi.org/10.1007/978-3-319-59250-3_14

Correa, J.R., Cristi, A., Oosterwijk, T.: On the price of anarchy for flows over time. In: Karlin, A., Immorlica, N., Johari, R. (eds.) Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, June 24–28, pp. 559–577. ACM, New York (2019). https://doi.org/10.1145/3328526.3329593

Dafermos, S., Sparrow, F.: Traffic assignment problem for a general network. J. Res. Natl. Bur. Stand. Sect. B Math. Sci. 91 (1969)

Department of Transport, Tourism and Sport: The Costs of Congestion. An Analysis of the Greater Dublin Area. Technical report. Government of Ireland (2017)

Harks, T., Peis, B., Schmand, D., Tauer, B., Vargas Koch, L.: Competitive packet routing with priority lists. ACM Trans. Eco. Comput. **6**(1), 4:1–4:26 (2018). https://doi.org/10.1145/3184137

Hoefer, M., Mirrokni, V.S., Röglin, H., Teng, S.: Competitive routing over time. Theor. Comput. Sci. **412**(39), 5420–5432 (2011). https://doi.org/10.1016/j.tcs.2011.05.055

Horni, A., Nagel, K., Axhausen, K. (eds.): Multi-Agent Transport Simulation MATSim. Ubiquity Press, London (2016). https://doi.org/10.5334/baw

Ismaili, A.: Routing games over time with FIFO policy. In: Devanur, N.R., Lu, P. (eds.) Web and Internet Economics, pp. 266–280. Springer International Publishing, Cham (2017)

Israel, J., Sering, L.: The impact of spillback on the price of anarchy for flows over time. In: Harks, T., Klimm, M. (eds.) Proceedings of the Algorithmic Game Theory – 13th International Symposium, SAGT 2020, Augsburg, September 16–18. Lecture Notes in Computer Science, vol. 12283, pp. 114–129. Springer, Berlin (2020). https://doi.org/10.1007/978-3-030-57980-7_8

Koch, R., Skutella, M.: Nash equilibria and the price of anarchy for flows over time. Theory Comput. Syst. **49**(1), 71–97 (2011). https://doi.org/10.1007/s00224-010-9299-y

Mobility and Transport: Transport in the European Union. Current Trends and Issues. Technical report. European Commission (2019)

Olivier, J., Peters, J.: Trends in Global CO_2 and Total Greenhouse Gas Emissions. PBL Netherlands Environmental Assessment Agency, The Hague (2020)

Peis, B., Tauer, B., Timmermans, V., Vargas Koch, L.: Oligopolistic competitive packet routing. In: Borndörfer, R., Storandt, S. (eds.) 18th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, ATMOS 2018, August 23–24, Helsinki. OASICS, vol. 65, pp. 13:1–13:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Wadern (2018). https://doi.org/10.4230/OASIcs.ATMOS.2018.13

Scarsini, M., Schröder, M., Tomala, T.: Dynamic atomic congestion games with seasonal flows. Oper. Res. **66**(2), 327–339 (2018). https://doi.org/10.1287/opre.2017.1683

Scheffler, R., Strehler, M., Vargas Koch, L.: Equilibria in routing games with edge priorities. In: Christodoulou, G., Harks, T. (eds.) Proceedings of the Web and Internet Economics – 14th International Conference, WINE 2018, Oxford, December 15–17. Lecture Notes in Computer Science, vol. 11316, pp. 408–422. Springer, Berlin (2018). https://doi.org/10.1007/978-3-030-04612-5_27

Sering, L.: Nash flows over time. Doctoral Thesis, Technische Universität Berlin, Berlin (2020). https://doi.org/10.14279/depositonce-10640

Sering, L., Vargas Koch, L.: Nash flows over time with spillback. In: Chan, T.M. (ed.) Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, January 6–9, pp. 935–945. SIAM, Philadelphia (2019). https://doi.org/10.1137/1.9781611975482.57

Sering, L., Zimmer, M.: Nash Flow Computation (2020). https://www.coga.tu-berlin.de/v_menue/download_media/nash_flow_computation/

Sering, L., Vargas Koch, L., Ziemke, T.: Convergence of a packet routing model to flows over time. CoRR abs/2105.13202 (2021). https://arxiv.org/abs/2105.13202

Tauer, B., Fischer, D., Fuchs, J., Vargas Koch, L., Zieger, S.: Waiting for trains: complexity results. In: Conference on Algorithms and Discrete Applied Mathematics, pp. 282–303. Springer, Berlin (2020)

Vickrey, W.S.: Congestion theory and transport investment. Am. Econ. Rev. **59**(2), 251–260 (1969)

Wang, Y., Szeto, W., Han, K., Friesz, T.L.: Dynamic traffic assignment: a review of the methodological advances for environmentally sustainable road transportation applications. Transp. Res. B Methodol. **111**, 370–394 (2018). https://doi.org/10.1016/j.trb.2018.03.011. https://www.sciencedirect.com/science/article/pii/S0191261517308056

Werth, T., Holzhauser, M., Krumke, S.O.: Atomic routing in a deterministic queuing model. Oper. Res. Perspect. **1**(1), 18–41 (2014)

Ziemke, T., Sering, L., Vargas Koch, L., Zimmer, M., Nagel, K., Skutella, M.: Flows over time as continuous limits of packet-based network simulations. Transp. Res. Proc. **52**, 123–130 (2021). 23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16–18 September 2020, Paphos. https://doi.org/10.1016/j.trpro.2021.01.014

Part II
Digitalization and Logistics

Intelligent Agents for Social and Learning Logistics Systems

Otthein Herzog and Ingo J. Timm

Abstract The digitalization of logistics processes is often based on distributed models and decentralized control. As these logistics models constitute an important part of Industrie 4.0 concepts they must be powerful enough to cover dynamic processes and must enable a host of functions such as goal-oriented, reactive, pro-active, communicative, cooperative, competitive, and learning behaviors. In addition, these distributed models must allow for simulating, planning, allocating, scheduling, and optimizing logistics tasks. This implies that they must be able to act through communication channels with each other thus establishing logistics social communities.

Multiagent Systems (MAS) have been around for more than 30 years and lend themselves to the implementation of these distributed models needed for autonomous and cooperating logistics processes. It will be described and also demonstrated by three case studies why MAS are well suited for social and learning logistics systems. It will be shown how the resulting distributed MAS models provide the required functionalities for production and transportation logistics including the handling of dynamic local events as an essential feature for the successful planning, scheduling, optimizing, monitoring, and control of global logistics processes.

Ingo J. Timm was affiliated to LogDynamics until 2006 as Postdoctoral Researcher for Artificial Intelligence at the Center for Computing Technologies of the University of Bremen, Germany

---

O. Herzog (✉)
Tongji University, Shanghai, People's Republic of China

University of Bremen, Bremen, Germany
e-mail: otthein.herzog@uni-bremen.de

I. J. Timm
SDS Branch Trier, Cognitive Social Simulation, German Research Center for Artificial Intelligence (DFKI), Trier, Germany

FB4 – Business Informatics I, Trier University, Trier, Germany
e-mail: ingo.timm@dfki.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_5

1 Introduction

Logistics as an interdisciplinary field of research between engineering (mechanical engineering, production technology), natural sciences (mathematics, computer science), and economics has gained considerable importance since the early 1990s. At the beginning of LogDynamics in Bremen, logistics was strongly influenced by economic and political changes, i.e., reduction of manufacturing depth within an enterprise, concentration on core competences, global production, as well as liberalization of European's transportation market. Production and transportation processes are getting more complex, dependencies between processes in logistics networks and the global distributed Production are increasing, must be coordinated throughout complex networks bringing together various enterprises. In consequence, complex planning, coordination, and optimization problems challenge information systems in logistics since the 90s. This development is followed by the trend to digitalization, i.e., the dramatic increase of electronic business and electronic commerce as well as the development of ubiquitous and mobile computing and the Internet of Things. Nowadays almost any device, product, or process step can be supported or controlled by digital means. These developments are leading to new and innovative products and processes, e.g., based on small lot sizes, customer demand, flexible response, and customized mass production (mass customization). In its beginning, these developments have been addressed to by the Collaborative Research Center on Autonomous Logistics (SFB 637, funded by the DFG[1]) in Bremen, Germany.[2]

A holistic perspective on logistics is widely proposed, which considers both the business strategy and technical execution of automated systems. Scholz-Reiter et al. (2004), for example, introduce a 3-layer architecture for the specification and management of complex logistics processes, consisting of the decision layer, the information layer, and the execution layer (cf. Fig. 1a). These systems are each assigned to a specific disciplinary focus: Organization and Management (Decision System), Informatics Methods and Information and Communication Technologies (I&C Technologies; Information System), as well as Material Flow and Logistics (Execution System). In the early days of logistics information systems, these three layers have been loosely coupled prohibiting situational or reactive behavior and decision-making, as the decision layer had no current information on the state or events in the execution system and computing resources in the information system where too limited for online (re-) planning. Developments of the Internet of Things are bridging the gap between information and execution system while agent technology mainly contributed to the information and decision system.

In context of Industrie 4.0, technological achievements should be used for real-time coupling of these three levels by bringing together Internet of Things and agent

---

[1] DFG: German Research Foundation (Deutsche Forschungsgemeinschaft)

[2] http://sfb637.uni-bremen.de/?&L=2

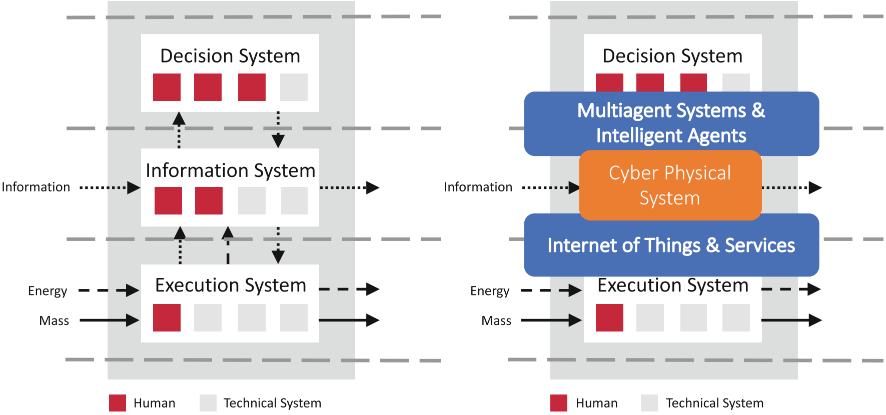


Fig. 1 Holistic perspective on logistics (**a**) with relevant technological innovations (**b**). adapted from (Scholz-Reiter et al. 2004)

technologies in the so-called Cyber Physical Systems (cf. Fig. 1b). Logistics systems are considered as autonomous (sub-) systems, where the behavior of individual actors depends on other actors or (sub-) systems in the "neighborhood." The autonomous systems interact with each other for pursuing their individual goals and the goals of the respective stakeholders. Thus, they are implementing local optimization, which leads to emergent behavior. As of this, it is uncertain how to guarantee, that individual behavior and emergent effects do not lead to global chaos, where no reliable prognosis on the system's outcome is feasible. Consequently, new challenges in modeling, engineering, and implementing autonomous logistics arise, which are related to behavior and management of groups in social sciences. In addition to the coordination problem of the 1990s, control and configuration of products and processes in production and logistics have also become increasingly relevant for future logistics information systems. To enable full potential of these technical and economic developments, future logistics information systems must handle knowledge about products and processes as well as their configuration or variation, additionally. As the requirements are shifting during runtime, future logistics information systems will require sophisticated social and learning abilities too.

In this chapter, we introduce intelligent software agents as promising technological solution for planning, designing, optimizing, and implementing social and learning logistics systems. We discuss innovative and application-oriented technological contributions from our research in LogDynamics on Intelligent Agents and modern logistics systems. Subsequently, we present case studies showing how Intelligent Agents implement logistics not only processes but outperformance conventional approaches. Complexity of logistics management is even of increasing complexity due to a wide variety of products and transportation processes, logistics networks, stakeholders, as well as domains.

2 Foundations of Multiagent Systems in Logistics

In the late 1980s, Intelligent Agents and Multiagent Systems (MAS) have been invented as innovative software technology in the field of distributed artificial intelligence for implementing distributed systems. As they do not require fully standardized communication or interaction mechanisms, MAS were seen as a silver bullet for highly heterogeneous real-world systems like logistics. Heinz Jürgen Müller (1997) introduced an engineering method for modeling MAS and identified three main characteristics that an application domain requires for a fruitful application of MAS: *"The application should show* ***natural distributivity****, e.g., autonomous entities, geographical distribution, distributed data; have a need for* ***flexible interaction****, e.g., there is no a priori assignment of tasks to actors, there are no fixed processes; be embedded in a* ***dynamic environment****, e.g. our physical world, artificial worlds like the internet, the world of finance" (Müller 1997, p. 218).* With the AWIC-methodology, HJ Müller also defines the core elements to be specified when applying MAS: agents, world, interaction, and coordination. The active elements in the system to be modeled should be represented by agents, the real-world is abstracted, such that the actions of agents can be performed, and relevant information can be perceived in the represented world. Interaction of and between agents as well as coordination are important mechanisms for the behavior of the group of agents. In consequence, design and implementation of agent interaction is a core challenge for engineering autonomous systems balancing restrictive interaction for enabling reliable behavior and permissive interaction allowing for flexibility and emergent behavior (Krempels et al. 2006). Furthermore, interaction between agents can dynamically constitute organizational structures, as required in business applications, i.e., the organization of MAS emerges from interaction (Ferber 1999; Timm et al. 2006a).

Knirsch and Timm (1999) discuss the applicability of MAS to logistics with respect to Müller's characteristics natural distributivity, flexible interaction, and dynamic environment. Moreover, MAS has the potential to influence the organization and configuration of businesses models themselves, especially concerning inter-business transactions. So, with MAS even short-term cooperation, e.g., temporal logistics networks, becomes feasible from an information systems perspective, as "Agents do not only 'know' about the system's configuration but even 'notice' events occurring and react accordingly" (Knirsch and Timm 1999, p. 214).

However, MAS could not gain acceptance at that time because, on the one hand, theory and practice were far apart and there was a lack of standardization and efficient implementation. With the FIPA-Initiative (Foundation for Intelligent Physical Agents), the community introduced a platform for standardization of an abstract architecture, an agent communication language (FIPA-ACL) as well as a broad set of interaction protocols (Poslad and Charlton 2001). These standardization efforts were liberally accepted by the community enabling cross-border communication and coordination of agents and MAS from various developers and on multiple platforms. In the German Priority Research Program on Intelligent

Software Agents and Business Applications (SPP 1083, funded by the DFG), the authors initiated a working group on bringing together multiple MAS in production logistics, developed at different research teams throughout Germany. With the Agent.Enterprise approach it has been shown that collaboration of heterogeneous MAS in logistics becomes feasible on the basis of the FIPA standardization (Stockheim et al. 2004). Finally, FIPA has been transferred to an IEEE standard in 2005.[3]

In contrast to conventional software systems, Intelligent Agents and MAS should implement reactive and deliberative behavior, proactivity, and flexible interaction skills (Wooldridge and Jennings 1995). However, theoretical and methodological research mainly focuses on specific aspects of agent's and MAS' behavior leading to a clustered world of formalizations, reasoning approaches, and frameworks and tools. Therefore, Timm developed the Discourse Agent architecture which brings together the formal specifications of individual agents and MAS as well as the FIPA standardization (Timm 2004). The Discourse Agent specifies an architecture for agent behavior, knowledge representation, and inferences. It strictly separates internal and external behavior in a 3-layer-architecture, i.e., to separate communication with other agents (communicator level), interface to real-world entities or business software (executer level), and internal decision-making for reasoning on the agent's next actions (controller level).

Next to communication behavior and social abilities, autonomy is one of the main features for Intelligent Agents and MAS in logistics. For a better understanding and characterization of MAS approaches in logistics, Timm (2006) introduced levels of autonomy ranging from strong regulation to operational, tactical, and strategic autonomy (see Fig. 2).

Strongly regulated systems are conventional software systems, which implement deterministic system's behavior, i.e., routines and behaviors are defined in design time and there is no functionality within the software system to adapt its behavior to a new or unforeseen situation. Low-level learning or reactive behavior, which can be adapted by the agent itself leads to operational autonomy. In operational autonomy the agents adapt action parameters, e.g., low-level artificial neural networks in context of robotics, or choose between equivalent action steps with respect to the current situation. If the agent is also adapting the action sequences themselves or its planning behavior, it implements tactical autonomy. Here, adaptation takes place between different courses of action, e.g., structural adaptation of hidden Markov models. On a strategic level, an agent is capable of altering the goal or desire set itself or adapts the evaluation of goals and desires with respect to midterm experience.

---

[3] http://www.fipa.org/

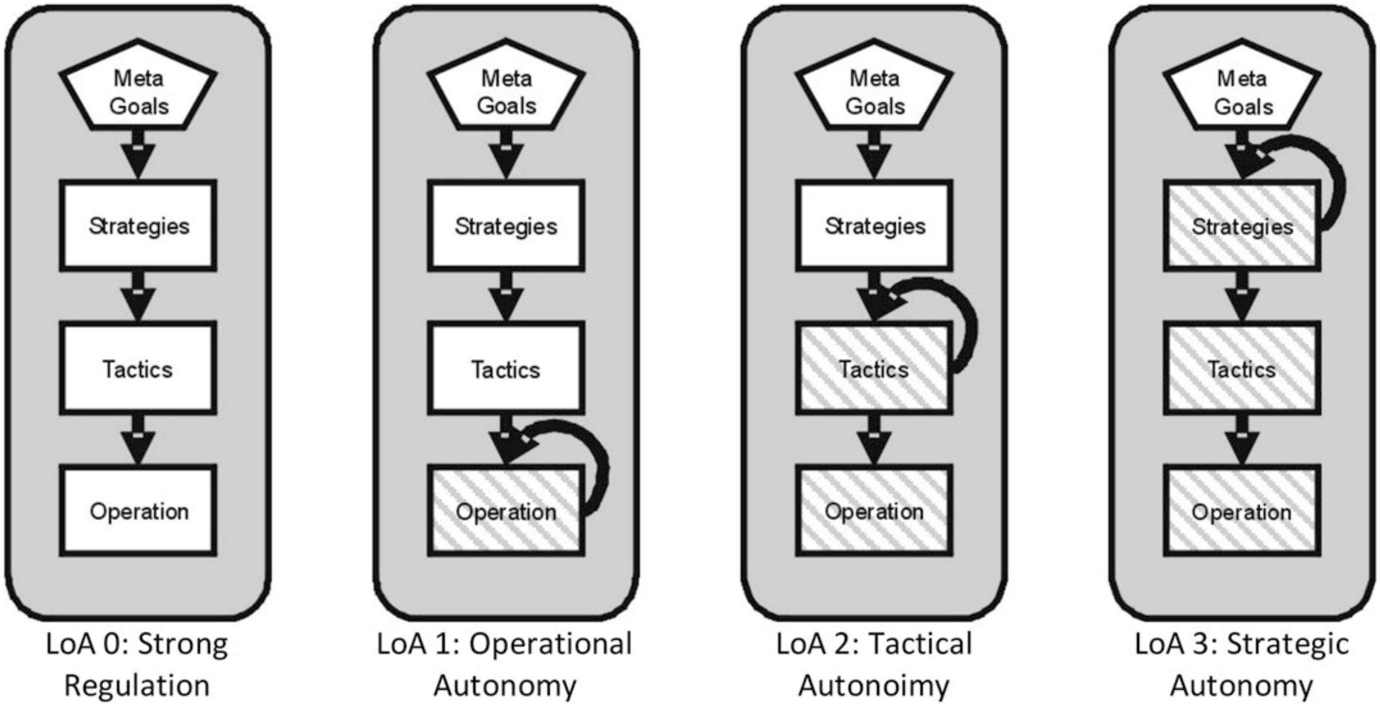


Fig. 2 Levels of Autonomy (Timm 2006)

3 Advanced Concepts for Multiagent Systems in Logistics

A distinctive feature of MAS is the potential for collaboration among agents that depends, however, on advanced agent features such as internal knowledge management, the ability to learn, to plan, to negotiate, and to communicate effectively among themselves in order to reach a common goal. These functionalities require in-depth knowledge about an application area (definitely not in a broad sense) and lead to autonomous MAS with potentially different operative roles, e.g., to select the best choice for the management of resources. The goal of such an agent society is rational and informed decision-making subject to their specific roles even if the partially visible environment of agents in an MAS increases the uncertainty in choosing the "right" decisions and the assessment of their utilities.

The most important prerequisite for informed decision-making is the ability to retrieve adequate knowledge for the assessment of situations that must be driven by utility, and also empirical knowledge for situation assessment (Gehrke and Wojtusiak 2008). This means that agents in an MAS must assume multiple roles: on the one hand, their specific domain-based roles, and, on the other hand, a knowledge management role where both producers and consumers are software agents. In order to achieve this, static domain knowledge must be available to the agents, e.g. a basic application domain ontology. In addition, empirical knowledge about the dynamics of the agent environment is needed that might be available through dedicated roles in order to enable situation assessment and prediction based on different knowledge models for decision-making involving decision, prediction, and classification models. Gehrke (2011) went a step further in researching how "relevant knowledge" could be secured by agents thus aiming directly at the semantic level and balanced decisions. A relevance-based information acquisition enables agents to identify, evaluate, and acquire exactly the information in a goal-oriented and efficient way that have the highest relevance for an impending decision. As there could be a large number of possibly relevant information pieces, Gehrke (2011) succeeded in developing a method to reduce the relevance assessment to a minimal number of information elements using especially the d-separation of nodes in decision networks as a criterion for the stochastic impact of information on the expected action utility. In this way, it was possible to exclude irrelevant and information of lesser value from further consideration.

Warden (2019) selected a machine learning and knowledge management approach for information retrieval in that he developed a comprehensive methodology comprising a set of knowledge management functions for individual information retrieval and knowledge creation. He focuses on knowledge creation through induction of classification models that enables some flexibility with regard to applicable knowledge creation. The ability for induced models constitutes a competitive advantage for a learning agent by incorporating relevant knowledge that is distributed among an agent community of practice to improve and shape individual induction. This enables a wide range of knowledge-based decision support systems where it is not necessary from the beginning on to define the complete knowledge necessary for an application. Instead, in a society of learning agents, it is possible for an agent "to ask around" for knowledge necessary for the next decision, e.g. some classification knowledge. In detail, it could be shown that multiagent interactive adaptation of individual classification models by a communication process between advisor and advisee agents can lead to knowledge creation for advisees. From the machine learning perspective, it was demonstrated that the technique of argument-based machine learning could be transferred from an agent-human setting into a setup of an agent society. As far as the technology has been evaluated, it could be confirmed that the modular architecture of a flexible integration of heterogeneous induction techniques by different machine learning frameworks was instrumental for these results.

While the agents in Warden (2019) communicated without any presupposition about the reaction of their counterparts, the work of Luhmann (1984, 1995) provides guidance on how to take also the "double contingency" into account that is supposed to build a foundation of a sociality in its way how communication partners make use of slight indications of mutual dependencies. Therefore, it is possible that a history about the communication can be created that leads to self-accelerating dynamics of social order generation. In this way, also a dynamic social order between agents in a multiagent setting can be established. Berndt (2015) could show that Self-Organizing Multiagent Negotiations allow for both competition and cooperation among concurrently acting agents using iterated negotiations as a basic interaction method where multiple agents compete for their individually best outcomes. However, by creating groups of agents with the same objective, there can be cooperation within such a group paired with competition with other agent groups which is limited because the standard negotiation methods among concurrently acting agents are not guaranteed to achieve efficient coordination results because they are not aware of the existence of their agent counterparts, their capabilities, their objectives, nor their current activities.

This leads to the observation that the participating agents are not able to decide which actions to select to achieve the coordination results.

Berndt (2015) could show that the method inspired by the Luhmann (1984, 1995) observations, the Self-Organizing Multiagent Negotiations generate social structures among the negotiating agents making use of adaptive coordination methods. These structures reduce an agent's decision-making contingency by making the results of possible actions expectable as its expectations are driven by its interactions and the agent will be able then to select its own actions according to the estimated responses by other agents. In this way, the agents can explore possible results of their potential actions and thus overcome lacking knowledge. The resulting self-organization avoids mutual disturbances of concurrent activities and allows the contributing agents to achieve their objectives by stabilizing their interactions. This anytime coordination algorithm can approximate a Nash equilibrium, and it was proven by Berndt (2015) that the asymptotic effort of Self-Organizing Multiagent Negotiations could be reduced by the goal-directed activity selections to $O(n^2)$ in the number of all participants in contrast to conventional negotiations with an asymptotic effort of O(n^3) at the expense of some additional memory requirements. The asymptotic computational effort could be shown to be O(n log n).

This social order generation process can be used to maximize the result quality even if the interaction and computation effort have to be limited, as it does not depend on the agent population size. A result quality evaluation could show that the capability to create a social order is crucial for multiagent teams or for the combinatorial coordination of multiple resources by a single individual agent and drastically outperforms conventional coordination methods. Using Self-Organizing Negotiations agent teams could be shown to outperform conventionally negotiating ones by up to 130% in their achieved welfare. This result is also well reflected in a real-world example where the Self-Organizing Multiagent Negotiations was applied to container logistics using sizable real-world data with almost 10.000 containers: the combined transportation and storage cost could be improved by almost 7%.

The results presented in this section underline the high potential of MAS in logistics. However, there is a great challenge in engineering specific systems by distributing roles and capabilities (Ferber 1999): Should every agent be capable and responsible for any tasks, or should each agent be specialized on one role only? The resulting problem is to dynamically balance efficiency (specialization) and reliability (generalization) with respect to current workload situations. Optimal solutions may be obtained by learning from human cooperative systems using psychologic research. Timm et al. analyzed intentional forgetting as an interdisciplinary concept to adapt human and AI-teams in the German Priority Research Program on Intentional Forgetting in Organizations (SPP 1921, funded by the DFG) (Timm et al. 2019). On this basis, an MAS extension for situational adaptation of processes and roles by learning and forgetting of actions, plans, and desires has been developed (Reuter et al. 2020; Timm et al. 2020). This leads to a new understanding of engineering and situational balancing specialization and generalization in logistics. Moreover, this approach can be used for hybrid teams, i.e., human-AI-workforces and has great potential for qualification planning of human teams and capability selection for MAS.

4 Case Studies

In this section, we present three case studies showing the potential of social and learning MAS in logistics. The first case study is dealing with production logistics. In manufacturing, efficiency and effectivity are often lost in process planning by fixing details on required production methods or the selection of resource types. Such pre-determinations prevent flexible reaction even of MAS to specific situations or events. In the manufacturing domain, the authors have developed an approach for integrated process planning and production control in a real-world scenario. On a larger scale, logistic networks can significantly benefit from the flexibility of MAS as shown in the SPP 1083 (Kirn et al. 2006). However, a more general approach is required, i.e., in the SFB 637, the concept of autonomous logistic entities has been developed. As in Industrie 4.0, such autonomous decision entities can be implemented using MAS. In various real-world applications, MAS have been proven to be beneficial in contrast to conventional approaches. Such applications and results are presented in the second case study. The third case study is addressed to the conceptualization of intelligent carriers, i.e., the intelligent container, which has also been developed within the SFB 637.

4.1 Case Study 1: MAS for Production Logistics

In the early days of production planning and process control, production and its manufacturing processes were primarily treated as coordination problems, i.e., processes were based on Tayloristic separation of preparatory activities (process planning) and implementation activities (production control) (Toenshoff 1999). The process planning was performed as a preparatory activity within or after product design. Production planning and control systems were restricted to optimize availability of material and resources as well as to allocate resources for large-scale or mass production. Except for the bill of materials, almost no information about the product itself was required in production planning. However, since the 1990s, the requirements to production logistics changed significantly due to trends like customized and personalized products, small lot sizes, or changing customer demand. For improved flexibility, pre-determination in preparatory steps like process planning should be reduced and the production planning and control system should be based on knowledge on the (customized/individualized) product, and on the skills and capabilities of the resources too.

In the IntaPS-projects (2000–2006) as part of the Priority Research Programme (SPP 1083) on Intelligent Agents and Business Applications, the authors developed an integrated approach to process planning and production control. The IntaPS approach consists of two components: an MAS implementing an electronic market as a decentralized planning component and formal representations of manufacturing features and skills and capabilities of resources. Agents are representing production orders or resources, like machine tools which are both representing digital twins of the real-world orders and resources responsible for negotiating resource allocation and scheduling production in a decentralized manner. For the technical implementation of such a system, the challenge arises, that there is almost no pragmatic architecture for implementing Intelligent Agents in such an environment. Many architectures developed for Intelligent Agents are based on unrealistic assumptions as consistency of goals. In logistics, it is well known that many standard objectives are competitive, e.g. maximizing utilization leads to high work-in-process and longer lead times.

Thus, Timm (2004) formally defined and specified a reference architecture, the so-called Discourse Agents. In contrast to other approaches, this architecture is based on potentially conflicting goals within an agent and between agents. The internal conflicts are handled by cobac (conflict-based agent control), which is an extension to the well-known belief-desires-intention architecture (Rao and Georgeff 1995; Wooldridge 2000). The cobac approach is based on a logical representation of goals and desires within an intelligent agent. As a first step in the deliberation cycle, an agent identifies accessible desires on basis of the current situation. In the next step, the desires are compared and evaluated due to relevance, chance, and risk of pursuing or ignoring available objectives. The desires are then processed pairwise, the potential for conflict or synergy is calculated, and a conflict resolution strategy is applied accordingly, e.g., merging desires, removing one of the two desires, etc. Additionally, Discourse Agents allow for dynamic adaptation and learning in negotiations using Markov chains with the OAC algorithm (2004). Both algorithms have been widely evaluated using agent-based simulation and an approach for testing autonomous systems has been developed (Timm et al. 2006b; Timm and Schumann 2009).

However, the strong focus on individual autonomy in the IntaPS project and the Discourse Agents led to limited efficiency on the shop floor level. Therefore, the question arises how the agents can consider objectives on the team level, e.g., of the shop floor at a whole or the enterprise, without limiting autonomy significantly. In an interdisciplinary cooperation, Timm and Hillebrandt (2006) have developed a framework for reflection in MAS inspired by Luhmann’s theory of group dynamics. The underlying idea here is that the agents act autonomously as long as global goals are not in danger. If global goals are expected to be missed, the different levels of autonomy will be restricted: first operational, then tactical, and finally strategic autonomy. Schumann et al. (2008) applied parts of this framework as “regulated autonomy” to the domain of job shop scheduling and evaluated the concept with respect to cost of control (switching between autonomous and central decision-making) and efficiency (lead time). The results show the feasibility as well as the efficiency of this approach.

In the IntaPS project as well as in the entire SPP 1083, it has been shown that Intelligent Agents and MAS are beneficial to incorporate flexibility in real-world applications as well as to coordinate logistics effectively (Kirn et al. 2006). However, in economics gain of flexibility is accompanied by loss of control on the strategic management level. Thus, a contradiction arises between strategic control of the logistics system and autonomous behavior of (sub-) systems (Dembski and Timm 2005). Consequently, additional measures must be taken to reach beneficial emergent behavior in autonomous logistics processes.

4.2 Case Study 2: MAS-Based Autonomous Logistics Processes

Besides the knowledge acquisition processes in MAS, a second important research area is how the planning, control, and monitoring of processes can be delegated to autonomous and decentralized agents in an MAS as the digital representatives of objects themselves. This approach is taken for the development of process control in Industrie 4.0 environments where supply network management becomes increasingly complex, dynamic, and distributed. The paradigm of autonomous logistics aims at automating process control by delegating decision-making to the participating logistics objects. Based on objectives imposed by their owners, these autonomous logistics entities can themselves plan and schedule their way through logistics networks. Each entity incorporates only its own parameters as well as those of cooperating entities. Therefore, the computational complexity can be reduced significantly and dynamics can be dealt with locally. Schuldt (2011) showed that agents in an MAS are well-suited to the implementation of decision-making by local data processing as well as the coordination of the individual autonomous logistics entities and therefore enable the application to real-world control of logistics processes where both logistics service providers and consumers can be modeled with intelligent agents. Schuldt (2011) developed specific interaction schemes that are required for autonomous logistics, e.g., interaction protocols for team formation of autonomous logistics entities that constitutes the foundation for jointly coordinating the primary logistics functions transport, handling, storage, and picking. In addition, interaction schemes for plan formation and team action were developed to satisfy complex logistics objectives.

Schuldt (2011) validated this novel approach based on real-world scheduling processes based on the procurement logistics processes of a major European retailer of consumer products with more than 1200 own shops and over 56,000 outlets in total. Previously manually controlled processes for the dispatching of over 11,500 shipping containers within one year were implemented MAS-based with autonomous control and validated with MAS-based simulation. The results show that automated process control for standard cases with autonomous logistics is applicable to satisfy even the challenging logistics demands of this retailer through allocating logistics resources efficiently and reliably. This approach even exceeds the efficiency of the manual approach: the MAS simulation showed savings of 2.6 million pallet-days in the warehouses per year by better utilizing free times at the container terminals.

Gath (2015) investigates how autonomous control can be implemented into current logistics processes for intelligent transport logistics, and which communication and negotiation mechanisms allow for MAS-based autonomous control in transportation logistics. He also specializes in the decision-making processes of autonomously acting agents to satisfy the dynamic requirements of transport service providers and their customers in order to identify the optimization potential of current transport processes in real-time situations.

In this context, Gath (2015) provides a framework to meet the advanced requirements for logistics transportation processes on complexity, dynamics, and customization, concentrating on stable communication and negotiation protocols for the synchronization in highly parallelized negotiations. In order to satisfy the real-time requirements, efficient routing algorithms were developed, which enable the agents to make optimal and also near-to-optimal decisions within a limited time frame where the concept of nested Monte Carlo search with policy adaptation was shown to be superior to a classical branch-and-bound approach to provide near-to-optimal solutions for routing problems. Through extensive evaluations of the agents' decision-making algorithms through established benchmarks, real-world groupage traffic, and real-world Courier, Express, and Parcel services (CEP), it could be shown that the developed algorithms outperform previous approaches. In contrast to full truckload traffic, in groupage traffic the complexity of process planning is even increased by highly volatile order situations, changing individual shipment qualities, and delivery/pickup time windows and further impeded by dynamics originating from unexpected events. The evaluation with real-world data provided by logistics enterprises show that multiagent-based autonomous control meets the sophisticated requirements of groupage traffic and CEP requirements and outperforms commercial dispatching software.

Icarte (2021a, b) extended this approach to the scheduling of equipment in open-pit mines that excel through ambitious production plans making use of expensive machinery equipment despite frequent dynamic events hampering the execution of the plans. He could show with real-world open-pit-mine data that his scheduling method implementing digital twins of the equipment by a decentralized MAS-based approach and emphasizing highly concurrent negotiations between the agents outperformed the current methods in respect to productivity, fleet size, and cost.

4.3 Case Study 3: Multiagent Systems and the IoT: The Intelligent Container

Hribernik et al. (2010) present a potential solution to integrate the information and material flows of autonomous cooperating logistics processes with a standards-based approach. An MAS models the information flow of autonomous logistics processes where the material flow is also represented using the appropriate logistics sensor data (e.g., location, speed, stops, temperature) and integrated with the information flow. Opening the decision horizon of an MAS in such a way allows for much more sophisticated information processing, inter-agent communication, and better decision-making. The proposed concepts for the connection of the information and material flows are based on the EPCglobal Standard Framework Architecture (EPCglobal 2007, 2008a, b, 2009).

An application scenario is shown in Fig. 3: There are different entity connection types that are part of the material flow, the communicating agents as its digital twins, and the relations among material and information flow. The integrative solution concept is based on four solutions to represent these relationships: (1) EPC for the unique identification of physical and virtual objects, (2) EPC Information Services as a gateway between MAS and the logistics objects in the material flow (see Fig. 3), (3) standard EPCIS events model the relations between the entities where adequate or proposed extensions to standard event types or new event types. (4) The integration of sensor data is achieved by an extension to the ALE (Application Level Events) standard. The application and extension of the widely adopted EPCglobal

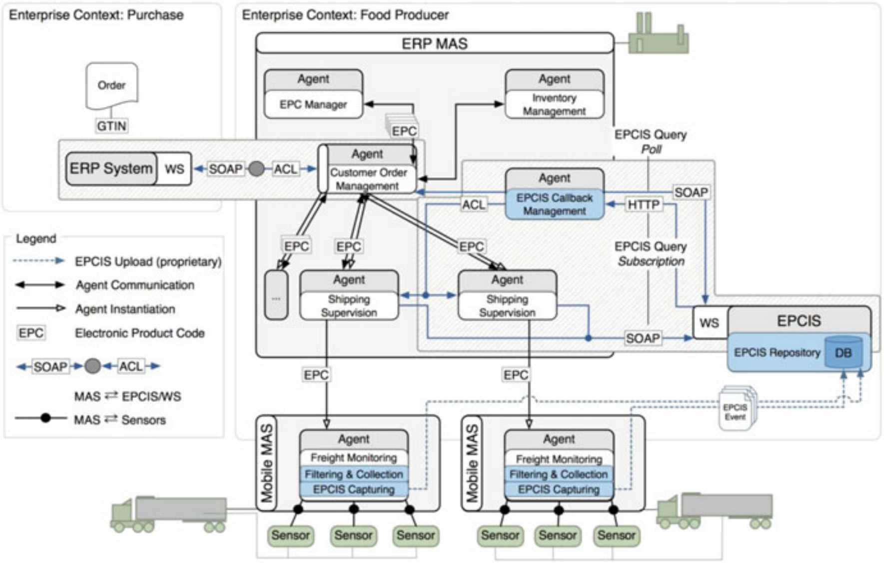


Fig. 3 Concept for an EPCIS-enabled implementation of the MAS outlined in the application scenario (Source: Hribernik et al. 2010, p. 60)

Architecture Framework to the problem allows for a standards-based solution that provides a platform-independent interoperability.

Dittmer et al. (2012) point out that the IoT focuses on technologies for data collection and distribution including standardization aspects such as EPCglobal (EPCglobal 2007, 2008a, b, 2009). However, this covers only the macro level of logistics systems whereas distributed logistics systems are designed for decentralized decision-making based on agent communication. Therefore, the approach of Autonomous Cooperating Agents defines the micro level of the logistic objects' behavior complementing the macro level. In the case of the Intelligent Container (Jedermann et al. 2007; Lang et al. 2014) this has been demonstrated by an implemented system for monitoring the ripeness of bananas in containers during their transportation from South America to Europe: here the Integrated Container can provide information about the shelf life of its cargo to the logistics network. To estimate the remaining shelf life from the data, decision support algorithms must be implemented based on the temperature history for the estimation of the changes of the fruits. The initial status of the fruits is determined by a visual analysis system of their color before they are loaded into a container. The container itself is equipped with a wireless sensor network that monitors the internal container status and also itself for the management of the sensor nodes and for failure detection. In addition, a decision algorithm has been developed that runs on sensor nodes and determines the temperature related quality losses, measurement intervals, sensor self-tests, and a sensor for the ripening indicator ethylene. Using the collected data during the transportation, a decision algorithm runs on sensor nodes that determines the remaining shelf life of the fruits. In this way, the Intelligent Container implements the new logistic paradigm of dynamic FEFO (First Expire First Out) as the remaining life time of the transported fruits is used to control the logistics process.

5 Discussion and Perspectives

It could be shown that the Distributed Artificial Intelligence technologies of Intelligent Agents and MAS is very well suited to build comprehensive models for autonomous cooperating logistics processes. This application area is characterized by highly uncertain events and processes where it is essential to rely on local resolutions that are worked out through the social interaction of the locally affected agents. In this way, decisions can be taken by design at the latest possible point in time, thus enabling the best possible outcome in a robust, resilient setting. In addition, learning (and forgetting) MAS models of logistics systems is able to preempt future adverse developments and thereby improve the local decision-making. The Collaborative Research Center on Autonomous Logistics (SFB 637) at the University of Bremen could demonstrate with its numerous publications and also by a substantial amount of successful SFB 637 start-ups that this route to the digitalization of logistics processes is a valid one and is being taken up by the logistics industry. The need for this kind of logistics digitalization became obvious during the first months of the COVID-19 pandemic where non-digitalized logistics networks had an almost complete breakdown, whereas digitalized logistics networks could recover within a comparably short time frame.

References

Berndt, J.O.: Self-Organizing Multiagent Negotiations Cooperation and Competition of Concurrently Acting Agents with Limited Knowledge. Doctoral Dissertation, University of Bremen (2015)

Dembski, N., Timm, I.J.: Contradictions between strategic management and operational decision-making – impacts of autonomous processes to decision-making in logistics. In: Palwar, K.S. et al. (eds.): Innovations in Global Supply Chain Networks. Proceedings of the 10th International Symposium on Logistics (ISL 2005), Lissabon, pp. 650–655 (2005)

Dittmer, P., Veigt, M., Scholz-Reiter, B., Heidmann, N., Steffen, P.: The intelligent container as a part of the internet of things. In: Proceedings of the 2012 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, pp. 209–214. IEEE, Piscataway (2012)

EPCglobal Inc: EPC Information Services (EPCIS) Version 1.0.1 Specification. EPCglobal Inc., Standard Specification, New Jersey (2007)

EPCglobal Inc: EPCglobal Object Name Service (ONS) 1.0.1. EPCglobal Inc., Standard Specification, New Jersey (2008a)

EPCglobal Inc: EPC Tag Data Standards Version 1.4. Standard Specification, New Jersey, EPCGlobal Inc (2008b)

EPCglobal Inc: The EPCglobal Architecture Framework, 1.3, Standard Specification. EPCGlobal Inc, New Jersey (2009)

Ferber, J.: Multi-agent systems: An Introduction to Distributed Artificial Intelligence. Addison-Wesley, Boston, MA (1999)

Gath, M.: Optimizing Transport Logistics Processes with Multiagent-based Planning and Control. Doctoral Dissertation, University of Bremen (2015)

Gehrke, J.D.:Relevanzbasierte Informationsbeschaffung für die informierte Entscheidungsfindung intelligenter Agenten. Doctoral Dissertation, University of Bremen (2011)

Gehrke, J.D., Wojtusiak, J.: Traffic prediction for agent route planning. In: Bubak, M., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) 8th International Conference on Computational Science 2008, vol. 3, pp. 692–701. Springer, Kraków, Poland (2008)

Hribernik, K., Warden, T., Thoben, K.-D., Herzog, O.: An internet of things for transport logistics – an approach to connecting the information and material flows in autonomous cooperating logistics processes. In: Proceedings of the 12th International MITIP Conference on Information Technology & Innovation Processes of the Enterprises, pp. 54–67 (2010)

Icarte, G.: Negotiations for Agent-based Concurrent Process Scheduling of Resources in Highly Dynamic Environments. Doctoral Dissertation, University of Bremen (2021a)

Icarte, G., Pinto, J.D., Herzog, O.: A dynamic scheduling multiagent system for truck dispatching in open-pit mines. In: Rocha, A.P. et al. (eds.): ICAART 2020, LNAI 12613, pp. 132–148 (2021b)

Jedermann, R., Gehrke, J.D., Becker, M., Behrens, C., Kluge, E.M., Herzog, O., Lang, W.: Transport scenario for the intelligent container. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation & Control in Logistics, pp. 393–404. Springer, Berlin (2007)

Kirn, S., Herzog, O., Lockemann, P., Spaniol, O. (eds.): Multiagent Engineering – Theory and Application in Enterprises. Springer, Heidelberg (2006)

Knirsch, P., Timm, I.J.: Adaptive multiagent systems applied on temporal logistics networks. In: Muffatto, M., Pawar, K.S. (eds.): Logistics in the Information Age. Proceedings of the 4th International Symposium on Logistics (ISL-99), Florence, Italy, pp. 213–218 (1999)

Krempels, K.-H., Spaniol, O., Scholz, T., Timm, I.J., Herzog, O.: Interaction design. In: Kirn, S., Herzog, O., Lockemann, P., Spaniol, O. (eds.) Multiagent Engineering – Theory and Application in Enterprises, pp. 383–403. Springer, Heidelberg (2006)

Lang, W., Janßen, S., Jedermann, R.: The intelligent container – a cognitive sensor net for fruit logistics. In: Proceedings of the 3rd International Conference on Sensor Networks (SENSORNETS-2014), pp. 351–359 (2014)

Luhmann, N.: Soziale Systeme. Grundriss einer allgemeinen Theorie. Suhrkamp, Frankfurt, Germany (1984)

Luhmann, N.: Social Systems. Translation of Soziale Systeme. Stanford University Press, Stanford, USA (1995)

Müller, H.-J.: Towards agent systems engineering. Data Knowl. Eng. **23**(3), 217–245 (1997)

Multiagent Coordination Enabling Autonomous Logistics. Doctoral Dissertation, University of Bremen (2010)

Poslad, S., Charlton P.: Standardizing agent interoperability: the FIPA approach. In: Luck, M., Mařík, V., Štěpánková, O., Trappl, R. (eds.): Multi-Agent Systems and Applications: 9th ECCAI Advanced Course, ACAI 2001 and Agent Link's 3rd European Agent Systems Summer School, EASSS 2001, Prague, Czech Republic, July 2–13, 2001 Selected Tutorial Papers, pp. 98–117 (2001)

Rao, A.S., Georgeff, M.P.: BDI-agents: from theory to practice. In: Proceedings of the First International Conference on Multiagent Systems, pp. 312–319. AAAI-Press, San Francisco, CA (1995)

Reuter, L., Berndt, J.O., Timm, I.J.: Exploring the effects of role design on agent learning. In: Schmid, U., Klügl, F., Wolter, D. (eds.) KI 2020: Advances in Artificial Intelligence, pp. 190–203. Springer, Berlin (2020)

Scholz-Reiter, B., Windt, K., Freitag, M.: Autonomous logistic processes – New demands and first approaches. In: Proceedings of 37th CIRP International Seminar on Manufacturing Systems, pp. 357–362. Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest (2004)

Schuldt, A.: Multiagent coordination enabling autonomous logistics. Doctoral Dissertation, University of Bremen, Springer, Heidelberg (2011)

Schumann, R., Lattner, A.D., Timm, I.J.: Management-by-exception – a modern approach to managing self-organizing systems. Commun. SIWN. **4**, 168–172 (2008)

Stockheim T., Nimis J., Scholz T., Stehli M.: How to build a multi-multi -agent system – the agent. Enterprise approach. In: Proceedings of the Sixth International Conference on Enterprise Information Systems – Volume 4: ICEIS, pp. 364–371 (2004)

Timm, I.J.: Dynamisches Konfliktmanagement als Verhaltenssteuerung Intelligenter Agenten. Doctoral Dissertation, University of Bremen. DISKI 283. infix: Köln (2004)

Timm, I.J.: Strategic Management of Autonomous Software Systems – Overview Article. TZI-Technical Report 35/2006 and part of Habilitation Thesis. University of Bremen: Bremen (2006)

Timm, I.J., Hillebrandt, F.: Reflexion als sozialer Mechanismus zum strategischen Management autonomer Softwaresysteme. In: Schmitt, M., Michael, F., Hillebrandt, F. (eds.) Reflexive soziale Mechanismen. Von soziologischen Erklärungen zu sozionischen Modellen, pp. 255–288. VS Verlag, Wiesbaden (2006)

Timm, I.J., Schumann, R.: Performance measurement of multiagent systems: towards dependable MAS. In: Proceedings of the 2009 Spring Simulation Multi Conference ADS, BIS, MSE, MSEng, vol. 41, pp. 177–184. SCS, San Diego, CA (2009)

Timm, I.J., Scholz, T., Herzog, O., Krempels, K.-H., Spaniol, O.: From agents to multiagent systems. In: Kirn, S., Herzog, O., Lockemann, P., Spaniol, O. (eds.) Multiagent Engineering – Theory and Application in Enterprises, pp. 35–51. Springer, Heidelberg (2006a)

Timm, I.J., Scholz, T., Fürstenau, H.: From Testing to Theorem Proving. In: Kirn, S., Herzog, O., Lockemann, P., Spaniol, O. (eds.) Multiagent Engineering. International Handbooks on Information Systems. Springer, Berlin, Heidelberg (2006b)

Timm, I. J., Reuter, L., Berndt, J. O., Ulfert, A.-S., Ellwart, T., Antoni, C. H.: Analyzing the effects of role configuration in logistics processes using multiagent-based simulation: An interdisciplinary approach. In: 52nd Hawaii International Conference on System Sciences, HICSS 2019, Maui, Hawaii, USA, January 8–11, pp. 1–10) (2019)

Timm, I. J., Reuter, L., Berndt, J. O.: Role assignment adaptation: an intentional forgetting approach. In: 53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7–10, pp. 1–10. (2020)

Toenshoff, H.K.: Dynamic modelling of process planning knowledge based on a shop-floor model. In: Proceedings of the 32nd CIRP International Seminar on Manufacturing Systems, pp. 497–504. Leuven, Belgium (1999)

Warden, T.: Interactive Multiagent Adaptation of Individual Classification Models for Decision Support. Dissertation, University of Bremen (2019)

Wooldridge, M.J.: Reasoning about Rational Agents. The MIT Press, Cambridge, MA (2000)

Wooldridge, M., Jennings, N.R.: Intelligent agents: theory and practice. Knowl. Eng. Rev. **10**(2), 115–152 (1995)

Semantic Interoperability for Logistics and Beyond

Marco Franke, Karl Hribernik, and Klaus-Dieter Thoben

Abstract A Semantic Mediator was conceived in the CRC 637 (The Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes” focused on adaptive logistic processes including autonomous capabilities for the decentralised coordination of autonomous logistic objects in a heterarchical structure.) to tackle problems of interoperability of heterogeneous information sources in autonomous cooperating logistics processes. Since the conclusion of the CRC 637, the Semantic Mediator has been developed further and successfully transferred to interoperability problems in different domains, including Industry 4.0, the Internet of Things, and Product Lifecycle Management. This paper will introduce the Semantic Mediator and present examples of its successful application.

1 Introduction

Now, more than ever, logistics is a cornerstone of the world’s economy, and its smooth operation is vital to supplying consumers and businesses around the globe. Strain is put on logistics networks, for example by global just-in-time supply chains (Pisch 2020) and the atomization of product-related services by mass customization and mass personalization (Kumar 2007). Logistics processes are challenged by the planning and implementation of the transport and storage of goods, considering time, quality, and costs. Moreover, the growing popularity of online shopping and the desire of customers for transparency and same-day delivery options challenge

---

M. Franke · K. Hribernik
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany
e-mail: fma@biba.uni-bremen.de; hri@biba.uni-bremen.de

K.-D. Thoben (✉)
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

Faculty of Production Engineering, University of Bremen, Bremen, Germany
e-mail: tho@biba.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_6

logistics. For example the number of postal deliveries has increased to over 3.5 billion in 2018—this corresponds to almost 12 million postal deliveries per delivery day. The vast majority, around 84 percent, are parcels (Infografik: Fast 12 Millionen Sendungen pro Zustelltag 2019). In today's world, the pressure of these challenges on logistics networks has become even more obvious. Solutions are sought for a more resilient response to supply chain disruptions caused, for example by COVID-19, the automobile microchip and buildings material shortage, and the Suez Canal blockage.

Concepts of autonomy and self-organization to improve the resilience of logistics networks by enabling decentralized decision-making were investigated in the Collaborative Research Center 637 "Autonomous Cooperating Logistics Processes" at the University of Bremen (CRC 637)[1] and are just as relevant to address today's challenges to logistics. By distributing decision-making over autonomous logistics entities throughout a logistics system, an increase in the robustness, flexibility, adaptability, and reactivity of the overall system was expected (Freitag et al. 2004).

Intelligent logistics objects are a core concept in the realization of autonomous cooperating logistics systems, in which the logistics objects themselves are endowed with the capability to *"process information, to render and to execute decisions on their own"* (Böse and Windt 2007). This means that each physical logistics entity, like a truck, a container, or a package, needs to be equipped with a component capable of processing information, making decisions, and communicating these decisions to other entities in the system. That is, each physical logistics entity requires a digital representation.

These digital representations need access to the right information at the right time to be able to make autonomous decisions towards their logistics goals. They need to access the many different data sources of the logistics IT ecosystem on demand to get that information. Today's logistics IT ecosystem is characterized by countless communication protocols, standards, data formats and terminologies. This so-called heterogeneity is the main challenge of interoperability in logistics and one which needs to be solved for intelligent logistics objects to make autonomous decisions based on accurate and timely information.

A Semantic Mediator was proposed and developed in CRC 637 as an interoperability approach that can contribute to solving this problem. Subproject C2 "Data Integration" successfully demonstrated the viability of the approach for achieving the interoperability of intelligent logistics objects with relevant data sources in the logistics IT ecosystem.

However, the Semantic Mediator's journey did not stop there. In the past decade, similar problems of interoperability have been identified in several different research areas. The applicability of the Semantic Mediator has been successfully investigated in several of these.

---

[1] The Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes" focused on adaptive logistic processes including autonomous capabilities for the decentralised coordination of autonomous logistic objects in a heterarchical structure.

For example Industrie 4.0 introduced the concept of Cyber-Physical Systems (CPS) to manufacturing. As autonomous, cooperative elements interacting in a system, CPS share many of the characteristics of intelligent logistics objects. This includes the need to reliably access decision-relevant information from heterogeneous information sources throughout, for example a Smart Manufacturing environment.

Another interoperability problem with similar characteristics can be found in Closed-loop Product Lifecycle Management (Closed-loop PLM). Closed-loop PLM posits closing the loops of all information systems throughout the product lifecycle, bridging the information silos that were previously separated from each other. The aim is to optimize processes by providing access to information from processes previously available.

In this article, the background of the interoperability problem in autonomous cooperating logistics processes is recapitulated. The problem of data heterogeneity identified there is discussed and updated to capture significant developments of the past decade, such as Cyber-Physical Systems, the Internet of Things and Digital Twins. The concept of semantic mediation was introduced as a solution to that problem. An implementation of the concept is presented, the subsequent evolution of which is traced through the different use cases beyond logistics the Semantic Mediator was applied to over the past 10 years. The paper closes with conclusions and an outlook to future work.

2 Background: The Need for Interoperability in Autonomous Cooperating Logistics Systems

For autonomous cooperating logistics processes to be realized, intelligent logistics objects must be able to process the information required to make and execute decisions towards their logistics objectives. This means that the intelligent logistics objects need to be able to access the information relevant to their decisions at any time throughout the logistics processes. Consequently, the objects not only need to be able to communicate with each other, but they need to be integrated into the overall logistics ecosystem in such a way that they can interoperate with any data source required to fulfil their information needs. The following subsections take a closer look at the characteristics of these data sources and what data representation, exchange formats and standards are widespread in logistics that need to be considered to achieve interoperability with those sources. This section attempts to bring the results of investigations into these issues done in CRC 637 up to date by reviewing the major developments in IT of the past decade, such as Cyber-Physical Systems, Digital Twins, and the evolution of the Internet of Things.

2.1 An Updated Look at Data Sources in Logistics

The major data sources relevant for autonomous cooperating logistics processes were categorized in CRC 637 as (1) Logistics IT Systems, (2) Intelligent Logistics Objects, (3) Digital Counterparts and (4) Sensors and Actuators (Hribernik et al. 2010). Whilst the first and last of the four categories are still appropriate today, the second and third need to be revisited in the light of the development of these topics over the past 10 years. Table 1 shows a restructured list of relevant data sources along with their respective interfaces and standards based on (Hribernik et al. 2011), which has been updated to include Digital Twins, Asset Administration Shell (AAS), CPS, and developments in the field of the Internet of Things. The following subsections discuss the individual categories in more detail.

Table 1 Updated table of major relevant data sources

Data sources	Type(s)	Interface/standard
Logistics IT systems	General	EDIFACT EANCOM
		EANCOM XML
		ebXML
	SAP compliant	SAP RFC (remote function call)
	Other	REST JSON XML Bespoke proprietary
Digital counterparts	Digital twins Cyber-physical systems Intelligent products Software agents	RAMI 4.0 AAS, others RAMI 4.0 AAS Dialog, PMI, EPCIS, OSGi ACL (agent communication language) Agent proxies Dialog
	Other	Bespoke proprietary
Internet of things	Internet of things	OPC-UA OPC DA
		OPC XML DA
		MQTT
		PMI
	Java-based	OSGi
	OGC compliant	SensorML
	General	GDI ORiN API
	Other sensors	Proprietary formats
	Other actuators	Proprietary formats

2.1.1 Logistics IT Systems

With regards to the first category, the IT systems used in the logistics sector today remain as varied, complex, and heterogeneous as investigated in CRC 637. Many different systems from different vendors are in use, many with their own proprietary data representations and interface schemas. To name but a few examples, they include Enterprise Resource Planning (ERP), Warehouse Management Systems (WMS), Transport Management Systems (TMS), Supply Chain Management (SCM), and Disposition Systems (DSS). In addition to solutions developed by solution providers, proprietary in-house solutions are widespread.

Exchange formats are the primary factors driving standardization in the field of logistics and the main means of interoperability between the relevant IT systems. Specific EDIFACT subsets like FORTRAS or the interface languages of popular ERP (Enterprise Resource Management) systems are predominant examples (Ballnus 2000). In addition, REST services and micro-service architectures have established XML and JSON as de facto data formats for sharing data. Moreover, document-based databases have supported these data formats to reach the persistence layer in legacy systems. Thus, application scenarios related to Industry 4.0 and IoT need to support these data formats.

2.1.2 Digital Counterparts

The second category originally encompassed Intelligent Logistics Objects, which constituted uniquely and automatically identifiable physical logistics objects. Category three included the decision-making components of intelligent logistics objects, such as software agents or holons. Over the past decade, these concepts have been consolidated together with similar concepts like Smart and Intelligent Products, Closed-loop PLM and others into notions such as CPS (Cyber-Physical Systems) and Digital Twins. For that reason, both have been consolidated here into Digital Counterparts.

CPS build upon concepts such as autonomous control, holonic manufacturing systems, intelligent manufacturing, IoT and embedded systems and are thus closely related to intelligent logistics entities (Lee 2006; Monostori 2014). CPS are *"systems of systems of autonomous and cooperative elements connecting with each other in situation dependent ways, on and across all levels of production, from processes to machines up to production and logistics networks, enhancing decision-making processes in real-time, response to unforeseen conditions and evolution along time".* (Cardin 2019). Like autonomous control in logistics, CPS systems promise to improve the adaptability, scalability, resiliency, safety, and security of industrial systems (Muller 2017), by contributing to making decentralized decisions (Lu 2017).

Digital Twins can be understood as comprehensive digital representations of physical assets, comprising their design and configuration, state, and behaviour. The term Digital Twin is still, however, used inconsistently throughout literature, and comprehensive reference models are lacking (Kritzinger et al. 2018; Lu et al. 2020). The RAMI 4.0 reference architecture now promotes the AAS interface as a single point of entry to the digital representations of physical assets and has been put forward as the link to Digital Twins. While more work is required before this is widely accepted, AAS needs to be taken into consideration when looking at interoperability with digital representations of physical assets like logistics entities going forward (Tantik and Anderl 2017; Anderl et al. 2018). Even though Digital Twins show clear parallels to concepts such as software agents or holons, the concept of autonomy has not yet been investigated adequately in conjunction with Digital Twins.

2.1.3 The Internet of Things

The fourth category includes sensors and actuators not covered by the previous three categories. In the past decade, the success of the Internet of Things has led to a mushrooming of IoT standards, with many hundreds competing for primacy in the marketplace. Even though the most recent years have shown a trend towards consolidation of standards in the industry around OPC-UA and MQTT, the multitude of implemented proprietary standards have only made the problem of interoperability more difficult. These changes have been reflected in the list of relevant data sources shown in Table 1.

2.2 Summarizing the Interoperability Problem in Complex Logistics Systems

The distributed, heterogeneous IT ecosystems outlined above, which prevail in complex logistics systems, demand a unique approach to interoperability. The heterogeneity and distribution of the data sources as well as their syntactic and semantic representations, pose a significant challenge. In addition to their geographical distribution, the data sources are distributed across multiple stakeholders in the logistics value network. Nevertheless, autonomous control cannot be realized without providing access to the required information in a predictable and reliable manner, regardless of the formal, structural, and physical properties of the underlying individual systems, standards, and formats (Hans et al. 2008). Some constraints identified in the CRC 637 no longer pose as much of a challenge as they did a decade ago. Specifically, concerns about the limited availability of intelligent logistics objects are almost negligible in today's world of ubiquitous wireless internet connectivity. The imminent widespread deployment of 5G networks will make this issue even less problematic, with the possible exception of logistics objects moving through remote areas.

3 The Problem of Data Heterogeneity

This section takes a closer look at the problem of data heterogeneity in distributed IT ecosystems such as complex logistics networks. It will furthermore introduce semantic mediation as the solution approach proposed in CRC 637 to achieve interoperability of intelligent logistics objects under these conditions.

3.1 Heterogeneity Classification

Data is represented using data formats that define the data's syntax. The syntax defines how the information has to be represented. The syntactical elements define which semantics of an item of information can be represented. For example the goal is to store a family tree in a data format. A representation in the CSV format could store a person, children, and parents as columns. In this case, additional knowledge is required to interpret the columns as parents and children. In contrast, XML or JSON can represent the family tree structure natively. The example demonstrates that different data formats have different syntax and capabilities. The interoperable representation of an item of information or the transformation between the data formats needs to consider the different capabilities of these data formats. Data integration conflicts may arise due to heterogeneity. A data integration conflict occurs if an item of information is extracted from one data source and inserted into another data source (Goh 1996; Wache 2003). For example the transformation of the family tree from JSON into CSV results in schematic conflicts because the taxonomy cannot be represented natively. An overview of possible data integration conflicts is shown in Table 2.

3.2 Interpretation Levels

Semantic mediation focuses on the interoperable representation of heterogonous data, which stores tuples containing data and meta-information. Using tuples prevents a false interpretation of data by different stakeholders. Information can be defined as *"…data that has been given meaning by way of relational*

Table 2 Types of possible data integration conflicts (Goh 1996)

Schematic Conflicts	Semantic Conflicts	Intensional Conflicts
Data type	Naming	Domain
Labelling	Scaling	Integrity constraint
Aggregation	Confounding	
Generalization		

Fig. 1 JSON example

```
{
    “parcel” :{
       “to”:{
              “name” : “Franke” ,
              “street ”: “Muster str.”
       }
}
```

connection . . .” (Bellinger et al. 2011). An unambiguous interpretation of the information inside each data source is possible on different levels: lexical, syntactical, morphological, semantic, and pragmatic understanding (Ören et al. 2007). Interpretation on lexical, syntactic, and morphological levels of understanding allows the structure of the information to be recognized and relevant entities to be grouped together. However, the meaning remains unclear. Common data formats for data exchange in Industrie 4.0 and IoT include CSV, JSON and XML. These data formats do not include enough meta-information to enable the false-free interpretation for the semantic understanding. An example of a parcel’s address in JSON, including the data and its meta-information, is shown in Fig. 1.

The parcel’s address defines on the morphological understanding that an addressee is defined by two attributes. The meaning of these two attributes is unclear without any additional meta-information. The reader cannot know that the first attribute describes the surname of a person and that the string “str.” inside the value of the second attribute street is an acronym for the street. This meta-information is not contained in the data. While a human reader immediately correctly interprets the JSON attributes name and street, this semantic understanding needs to be added programmatically as part of a data integration solution.

4 Solution Approach: A Semantic Mediator for Complex Logistics Systems

The goal is to represent the information along with all required meta-information to facilitate interpretation on the level of semantic understanding, rendering the information independent of its original data format and neutralizing data format restrictions with respect to syntax and semantics. To do so, each information needs to be represented in the interoperability model supporting the appropriate amount of meta-information.

The proposed approach uses a network of ontologies as an interoperability model. Ontologies are formal, partial specifications of an agreement over the description of a domain (Guarino 1998). Specific domains are modelled since broader ontologies are generally not convincing or feasible (Masolo et al. 2002).

Ontologies are modelled for different levels with specific purposes. Ontologies that model basic concepts and properties are called upper ontologies and are important for integrating heterogeneous knowledge from different sources (Mascardi et al. 2007). Examples of upper ontologies include Basic Formal Ontology (BFO), Business Objects Reference Ontology (BORO), Conceptual Reference Model (CIDOC), and Descriptive Ontology for Linguistic or Cognitive Engineering (DOLCE). Domain ontologies make use of upper ontologies for basic concepts and properties but describe all specifics with their own concepts. Domain ontologies in logistics include GenCLOn for urban freight transport (Anand et al. 2012) or ontologies for logistics services (Scheuermann and Hoxha 2012; Deng et al. 2019).

Whilst the interoperability approach proposed here uses upper and domain ontologies, it does not use a fixed ontology network or monolithic ontology since the data sources it needs to support are varied and may change over time. It proposes a loose coupling of existing domain ontologies to create a network of ontologies based on data sources such as intelligent logistics objects, IoT, CPS, and legacy systems. The adaptive ontology network proposed is appropriate for the movement of objects throughout a complex network and across stakeholder and process boundaries. Moreover, the approach needs to support a bidirectional flow of information between systems while allowing each system to access the information in its own data format. To achieve this, the role of the ontology network is as an intermediate representation, which is shown in Fig. 2. Each system transforms its data into the interoperability model. Based on the interoperability model, the transformation is possible with one further transformation step into each target data format. This approach guarantees scalability with respect to the number of involved systems.

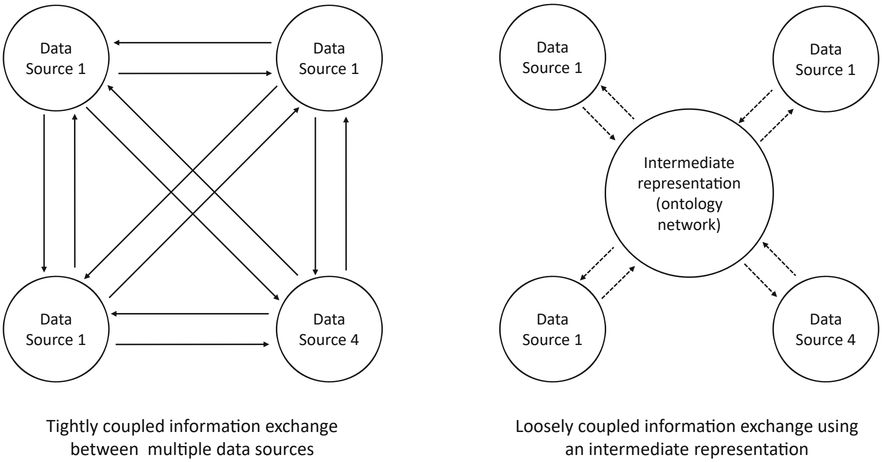


Fig. 2 Tight coupling vs. integration approach with an intermediate representation

4.1 Data Integration Approach

The information gathered from a data source can be mapped onto the interoperability model using Global as View (GAV), Local as View (LAV) or hybrid (GLAV) approaches (Lenzerini 2002). The approach proposed here uses GAV. A data source schema is mapped onto a global view, which is the interoperability model as a network of ontologies. Each information query in GAV needs to be formulated based on the global view based on ontological concepts and properties. This allows a query to be designed independently of specific data sources. The proposed approach uses the query language SPARQL to support this.

A virtual data integration approach, which aggregates the data on the fly and does not store the result, is followed. It is executed each time a SPARQL query is issued. The SPARQL query defines which amount of information should be requested from all connected heterogeneous data sources. The query-based data integration uses a mediator-based approach that collects and aggregates data from heterogeneous data sources. The approach, shown in Fig. 3, uses the component Semantic Mediator to aggregate partial results from the data sources. Each data source is connected to the Semantic Mediator via a wrapper, which translates data from the local view into the global view. In the following, both components are presented in detail.

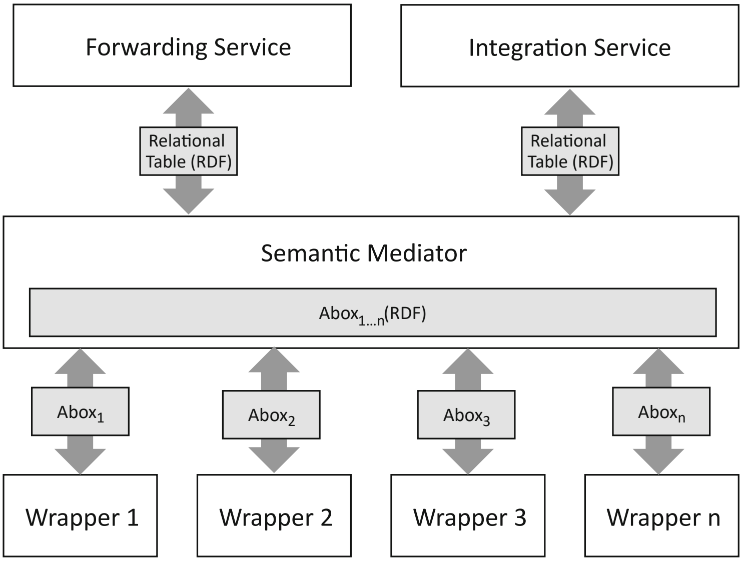


Fig. 3 The semantic mediator approach

4.1.1 Semantic Mediator Core Component

The core component implements Semantic Mediator functionality. Its input is a SPARQL query, and its output is an intermediate representation as ontology. The data integration starts with the identification of relevant data sources. The identification process maps the query's ontological concepts and properties on the local ontologies provided by the data sources. The Semantic Mediator can thus derive data source specific queries and issue them to the IT systems in question. These queries only include the ontological concepts that can be addressed by the data source. Each data source responds to the respective query by providing the Semantic Mediator with a set of individuals for the requested concepts. The semantic mediator collects all individuals and joins them. The resulting single ontology includes both the individuals and the definition of the concepts. This ontology can be delivered as a result or used as an intermediate representation for post-processing.

4.1.2 Wrapper

Wrappers implement the links between data sources and the Semantic Mediator. These links are based on configurations that map the local views of the data sources onto the global view of the Semantic Mediator. There is no need to implement interfaces to bind a data source to a wrapper or implement the data acquisition and the data transformation for each linked data source. Different kinds of wrappers are required to bind specific types of data sources. Wrappers have been prepared to connect to different relevant data sources listed in Table 1, such as EDIFACT EAONCOM, SQL databases, CSV files, XML files, REST services or streaming platforms such as Kafka. Each wrapper configuration contains an ontology defining the content of the data source and a wrapper type-specific mapping file defining the mapping between the source schema and the ontology schema.

5 Data Transformation Examples

The following example shows how a JSON Wrapper transforms JSON data into an RDF/XML ontology. Figure 4 shows the company name represented in different data formats. The company's transformation of all shown data source snippets would need the application of the EDIFACT, ANSI ASC X12, XML, and JSON wrappers. The outcome of these wrappers would be four individuals of the concept Company. The Semantic Mediator receives the four individuals and performs a join. In this example, all four data snippets contain a similar company name. Thus, the data integration would result in only one individual inside the ontology (see Fig. 6). The following shows the transformation of a JSON String into an ontological individual.

EDIFACT

```
NAD+BT+MyCompany::91‘
```

ANSI ASC X12

```
PER*CR*MyCompany*TE
```

O-MI

```
<omi:msg xmlns="odf.xsd"
xsi:schemaLoca-
tion="odf.xsd">
   <Objects>
       <Object>
             <id>My-
Company</id>
       </Object>
   </Objects>

  </omi:msg>
```

Custom

```
{
  "company":{
       "name": "MyCompany"
  }
}
```

Fig. 4 Name of a company represented in different data formats

```
<DatatypePropertyMapping propertyName=”Name”
Unit=”String”>
       <SubClass>Company</SubClass>
       <Tag>Root/company/name</Tag>
</DatatypePropertyMapping>
```

Fig. 5 Transformation rule for the property name

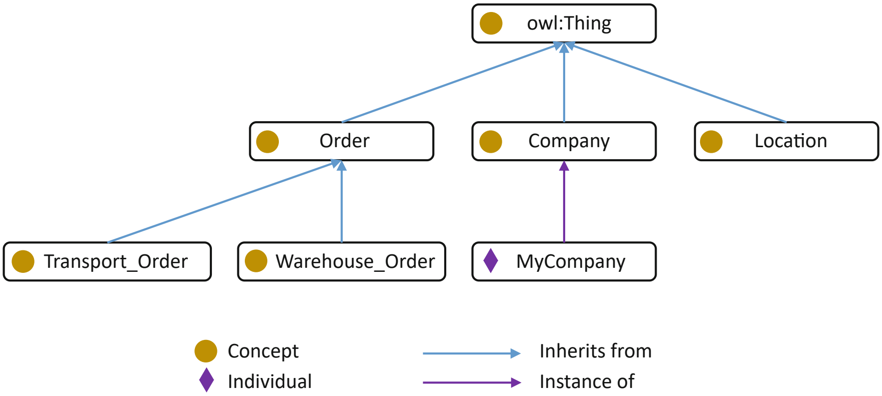


Fig. 6 The result of the transformation into an ontology

The JSON Wrapper follows the configured transformation rules and applies them to the data. Figure 5 shows a transformation rule that transforms the JSON String of the example in Fig. 4 into the ontology of Fig. 6.

For that purpose, the transformation rule defines that the rule is applicable for the property Name of the concept Company. Subsequently, the rule defines that the value for property Name is located in the path Root/company/name in the JSON file. Thus, the wrapper applies all rules sequentially, and the application of each rule extends the ontology with additional statements.

The presented example for the transformation of data into ontologies works similar for all wrappers. The only exception is that the structure of the transformation rules is dependant on the data source.

6 Evolution of Semantic Mediator

The semantic mediation approach for interoperability in complex logistics systems was proposed in 2009. Since then, the approach has been successfully applied to different interoperability problems in various industrial sectors. While the core design of the approach stood the test of time, and the overall system has evolved. The following describes a series of select application scenarios along with their impact on the mediator as a historical outline. Subsequently, the evolution steps are presented as a timeline.

6.1 Application Scenarios as a Historical Outline

In the subproject C2 of the Collaborative Research Centre SFB 637, research was carried out between 2008 and 2012 on developing an interoperability approach for the integration of logistics data to support self-controlling logistics processes. The main result of the work was a prototypical mediator component called "Semantic Mediator“. It focused on a virtual data integration approach whereby the logistics data sources were the primary data sources. This prototype was applied as preliminary work for later research projects.

6.2 Interoperability for Cyber-Physical Systems

The goal of the research project CyProS was the development of a representative spectrum of CPS modules for production and logistics systems. For that purpose, a framework for designing CPS-based solutions in production environments was in the focus.

The framework emphasized the need for interoperable information exchange for the seamless collaboration between CPS and legacy systems. The Semantic Mediator made it possible to transform data from the CPS and the legacy systems into one ontology to create a common global view (as motivated in Fig. 7) for

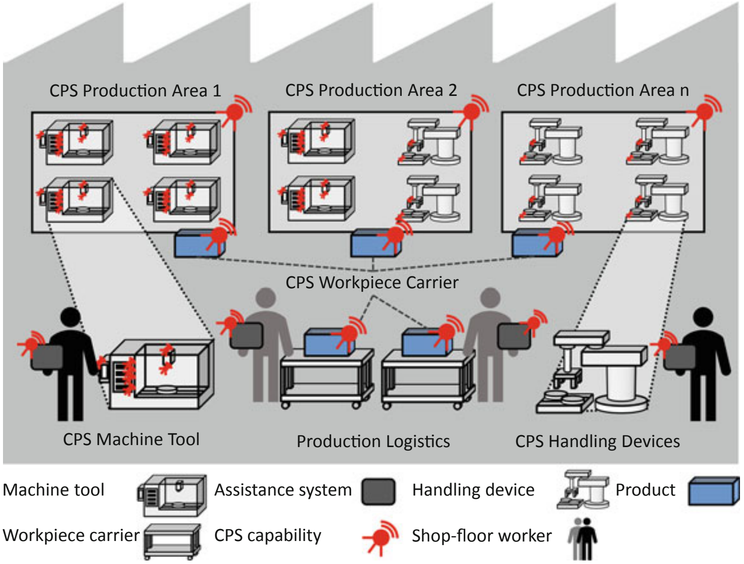


Fig. 7 Enabling interoperability between CPS and legacy systems (adapted from Reinhart et al. 2013)

autonomous and decentral decision-making processes. The dynamic registration and deregistration of data sources for an open-world application scenario was the focus. Here, the adaptation of the domain ontology in the role of a shared global view was developed depending on the connected systems. In addition, wrappers for CPS-specific interfaces were developed.

6.3 Interoperability of Product Usage Information for Product-Service-System Improvement and Design

A further interoperability problem was found in the field of Product-Service Systems (PSS) (Hribernik et al. 2018). These combinations of services and products are reliant on product usage information (PUI)—information about how a product is used—to help companies provide services offers for their products throughout their lifecycles (see Fig. 8). Previously, PUI was collected in processes such as customer relations management and MRO (maintenance, repair and overhaul) via repair logs, call centres or helpdesks. The digitalization of the product lifecycle, the Internet of Things, CPS, and social media have introduced new sources of valuable PUI which can be used in product and service development systems, e.g. Product Data Management (PDM), Product Lifecycle Management (PLM), Computer-Aided Technologies (CAx) and simulation tools, and the related processes and

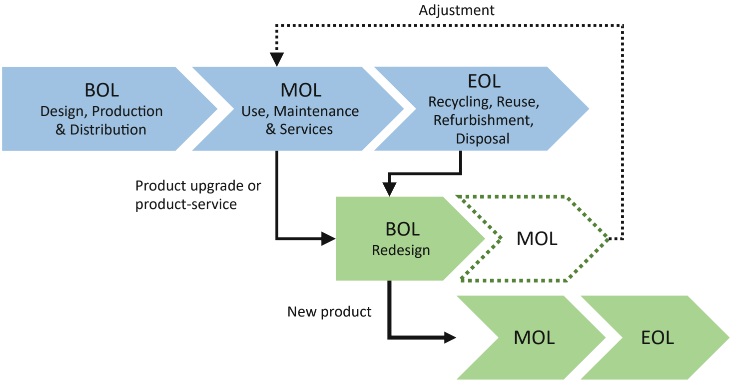


Fig. 8 PUI feedback loops in PSS Improvement and Design (adapted from Hribernik et al. 2016)

methodologies to improve current PSS and design new ones. The Semantic Mediator was consequently extended to efficiently and seamlessly integrate and apply PUI to PSS (re-)design processes. A major extension in this context was the development of a wrapper that uses text mining algorithms and NLP (natural language processing) to extract information from social media and integrate it with other PUI for use in PSS development and design.

6.4 Sustainable Manufacturing: Extending the Useful Life of Major Capital Investments and Large Industrial Equipment

A further use case for the Semantic Mediator was found in the field of sustainability in manufacturing, more specifically in extending the useful life of major capital investments and large industrial equipment. Here, a platform is being developed which offers services, ranging from the Digital Twin's setup, modernization actions to diagnose and predict the operation of physical assets to the refurbishment and remanufacturing activities towards End-of-Life (see Fig. 9). These services depend heavily on information gathered from machines and legacy systems. The interoperability challenge here is to achieve a scalable and extendable service platform. For that purpose, the Semantic Mediator enables the translation of the streaming data as well as legacy data via specific ontologies into the input data formats of the service layer. To support this feature, there was a paradigm shift in the data integration approach. Static data sources were until then the primary goal for the data integration, which stored the data permanently and can be requested on the user demand. This assumption changed to support also dynamic data sources whose data are volatile and must be transformed immediately independently from the current user demand. Examples of these new types of data sources are comments

Fig. 9 A platform for extending the useful life of major capital investments and large industrial equipment (The LEVEL-UP Project 2019)

in social media distributed by websites or sensor values distributed by distributed event streaming platforms such as Apache Kafka.

The results were new wrapper types, and the support of interoperable information flows based on streaming data. Moreover, the virtual data integration approach shifted to hybrid approach that combines the virtual and the physical data integration approaches. Current research projects suggest another paradigm shift. The focus of the ontology shifts from the target result to an intermediate representation. The uses cases need to consume the integrated information not as ontology but rather requested the translation into the target data formats of the use cases. For example sensor values should no longer be provided as an ontology but should be uploaded directly to an influx database for AI methods.

7 Outlook and Conclusion

The original aim of the Semantic Mediator was to solve the interoperability problem of intelligent logistics objects in autonomous cooperating logistics processes. The Semantic Mediator was successfully shown to be able to contribute to solving the problem with its virtual data integration approach, which focuses on providing exactly the right amount of information required on demand and defined by each individual query. The Semantic Mediator was shown to be able to provide intelligent logistics objects with the information required for taking autonomous logistics decisions. That information is stored in heterogeneous, highly distributed systems such as logistics enterprise systems, Internet of Things data sources, sensors, and embedded systems. It was shown that the chosen interoperability approach was capable of solving the heterogeneity conflicts that may arise in the heterogeneous and distributed IT ecosystem found in complex logistics systems. The resulting software solution, the Semantic Mediator, comprised the core mediator component along with a number of wrappers focused on the logistics application domain. These included EDIFACT EANCOM, CSV, and EPCIS wrappers, which were themselves configurations of generic wrappers developed to handle semi-structured texts, database management systems, and enterprise systems. In addition, a gateway module was developed to provide the Semantic Mediator with capabilities to access the Internet of Things and other embedded devices.

The subsequent research in complementary problem areas and industrial sectors led to an extension of the capabilities of the Semantic Mediator. These included wrappers for the connection of SQL and XML data sources which were added up to 2015. The next evolutionary step was taken in 2017 when data integration was expanded to support also physical data integration. For this purpose, the possibility was created to forward data to a data sink. A triple store was selected as a data sink for this purpose. Furthermore, a new generation of wrappers has been introduced that significantly improved the capabilities of the mediator to handle dynamic data sources, such as streaming sensors or social media data. The current expansion stage supports Kafka as a dynamic data source and provides the feature that transformed data can also be published on streaming platforms such as Kafka. This enables the mediator to homogenize and stream data from heterogeneous data sources using a network of ontologies within the framework of a continuous flow of information. Possible consumers of these streams are services such as predictive maintenance or anomaly detection, which rely on data from various sources.

The Semantic Mediator is well prepared for future digitization challenges beyond Industry 4.0, Autonomous Logistics, Digital Factories and Closed-loop PLM. The challenging requirements of interoperability for autonomous cooperating logistics processes on the one hand, and on the other the concept of intelligent logistics objects which foreshadowed later developments such as CPS, Digital Twins, Smart and Digital Factories influenced the design of the Semantic Mediator in such a way that it proved applicable to many novels and challenging interoperability problems, a selection of which are outlined above. It is a testimony to the research done in CRC 637 and more broadly in LogDynamics that even after 10 years, the principles underlying the Semantic Mediator can still be considered novel and relevant to today's interoperability problems.

Even though the Semantic Mediator has evolved significantly over the past decade, there remain research challenges that have not yet been solved. Apart from tapping additional application scenarios, the capability for the mediator to (semi-)automatically integrates new and unknown data sources has not yet been realized. This adaptive, (semi-)automatic approach to semantic data integration could significantly improve the flexibility of autonomous processes in logistics, manufacturing, the product lifecycle, and other domains. Starting points for research towards adaptive semantic data integration have already been identified and include applying principles of ontology learning, algorithmic ontology mapping, and methods of artificial intelligence and software engineering for automatic configuration and deployment. That means the next 10 years of research into interoperability in complex industrial systems will remain as exciting and productive as the last decade.

Acknowledgements This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes", the German Federal Ministry of Education and Research (BMBF) as a part of the project CYPROS (Grant Agreement No. 02PJ2477), and the European Commission through the projects FITMAN (Grant Agreement No. 604674), FALCON (Grant Agreement No. 636868) and LEVEL-UP (Grant Agreement No. 869991). The contents of this paper reflect only the authors' view, and the respective funding organizations are not responsible for any use that may be made of the information it contains. The authors wish to acknowledge the funding organizations and all project partners for the fruitful collaboration.

References

Anand, N., et al.: GenCLOn: an ontology for city logistics. Expert Syst. Appl. **39**(15), 11944–11960 (2012). https://doi.org/10.1016/j.eswa.2012.03.068

Anderl, R., et al.: Digital twin technology – an approach for Industrie 4.0 vertical and horizontal lifecycle integration. Inf. Technol. **60**(3), 125–132 (2018). https://doi.org/10.1515/itit-2017-0038. De Gruyter Oldenbourg

Ballnus, R.: Erfolg mit EDI und E-Commerce: Handlungsempfehlungen für die Abstimmung und Organisation interorganisationaler Netzwerke.. Available at: https://books.google.de/books?hl=en&lr=&id=7tQGL6hi19IC&oi=fnd&pg=PA11&dq=Erfolg+mit+EDI+und+E-commerce+-+Handlungsempfehlungen+f%C3%BCr+die+Abstimmung+und+Organisation+interorganisationaler+Netzwerke&ots=ecGT8N0-jo&sig=JNdAjWpAUQvkjxMVEuoTI_hPtCM (2000). Accessed 5 July 2021

Bellinger, G., Castro, D., Mills, A.: Data, Information, Knowledge, and Wisdom (2011)

Böse, F., Windt, K.: Catalogue of criteria for autonomous control in logistics. In: Understanding Autonomous Cooperation and Control in Logistics: the Impact of Autonomy on Management, Information, Communication and Material Flow, pp. 57–72. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-47450-0_5

Cardin, O.: Classification of cyber-physical production systems applications: proposition of an analysis framework. Comput. Ind. **104**, 11–21 (2019). https://doi.org/10.1016/j.compind.2018.10.002

Deng, Q., et al.: An ontology framework for multisided platform interoperability. In: Proceedings of the I-ESA Conferences. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-13693-2_36

Freitag, M., Herzog, O., Scholz-Reiter, B.: Selbststeuerung logistischer Prozesse–ein Paradigmenwechsel und seine Grenzen. Industrie Management, 20:23–27. Available at: http://www.sfb637.uni-bremen.de/pubdb/repository/SFB637-A5-04-002-IJ.pdf (2004). Accessed 5 July 2021

Goh, C. H.: Representing and reasoning about semantic conflicts in heterogeneous information systems. Massachusetts Institute of Technology. Available at: https://elibrary.ru/item.asp?id=5551531 (1996). Accessed 6 July 2021

Guarino, N.: Formal ontology and information systems. In: Formal Ontology in Information Systems, pp. 3–15. IOS Press, IOS Press, (1998). Available at: https://books.google.de/

books?hl=en&lr=&id=Wf5p3_fUxacC&oi=fnd&pg=PR5&dq=Guarino,+N.:+Formal+Ontology+in+Information+Systems&ots=noSG1StCNL&sig=OiecmvXsp1aw_6l42X2OaW7_kr4. Accessed 5 July 2021

Hans, C., Hribernik, K.A., Thoben, K.-D.: An approach for the integration of data within complex logistics systems. In: Dynamics in Logistics, pp. 381–390. Springer, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-76862-3_38

Hribernik, K.A., et al.: A semantic mediator for data integration in autonomous logistics processes. In: Enterprise Interoperability IV, pp. 157–167. Springer, London (2010). https://doi.org/10.1007/978-1-84996-257-5_15

Hribernik, K.A., et al.: Impacts of data integration approaches on the limitations of autonomous cooperating logistics processes. In: Autonomous Cooperation and Control in Logistics, pp. 247–267. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-19469-6_18

Hribernik, K.A., et al.: Feedback mechanisms across the lifecycle for customer-driven optimisation of innovative product-service design. In: Zelm, M., Doumeingts, G., Mendonca, J.P. (eds.) Enterprise Interoperability in the Digitized and Networked Factory of the Future (I-ESA 16 Workshop Proceedings). ISTE Ltd, Washington, DC (2016)

Hribernik, K. et al.: Towards a platform for integrating product usage information into innovative product-service design. In: 2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 – Proceedings, pp. 1407–1413 (2018). https://doi.org/10.1109/ICE.2017.8280047

Infografik: Fast 12 Millionen Sendungen pro Zustelltag.: Statista.. Available at: https://de.statista.com/infografik/9992/in-deutschland-von-den-paket-und-kurierdiensten-befoerderten-sendungen/ (2019). Accessed 6 July 2021

Kritzinger, W., et al.: Digital twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine. **51**(11), 1016–1022 (2018). https://doi.org/10.1016/j.ifacol.2018.08.474

Kumar, A.: From mass customization to mass personalization: a strategic transformation. Int. J. Flex. Manuf. Syst. **19**(4), 533–547 (2007). https://doi.org/10.1007/s10696-008-9048-6

Lee, E. A.: Cyber-physical systems – are computing foundations adequate? In: NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap. Austin, Texas, pp. 1–9. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.8011&rep=rep1&type=pdf (2006). Accessed 18 July 2019

Lenzerini, M.: Data integration. In Association for Computing Machinery (ACM), p. 233 (2002). https://doi.org/10.1145/543613.543644

Lu, Y.: Industry 4.0: a survey on technologies, applications and open research issues. J. Ind. Inf. Integr. **6**, 1–10 (2017). https://doi.org/10.1016/J.JII.2017.04.005. Elsevier

Lu, Y., et al.: Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. Robotics and Computer-Integrated Manufacturing. **61**, 101837 (2020). https://doi.org/10.1016/J.RCIM.2019.101837. Pergamon

Mascardi, V., Cordì, V., Rosso, P.: A comparison of upper ontologies. In: WOA 2007 – 8th AI*IA/TABOO Joint Workshop ‘From Objects to Agents’: Agents and Industry: Technological Applications of Software Agents, pp. 55–64. Available at: https://www.researchgate.net/profile/Viviana-Mascardi/publication/228674060_A_Comparison_of_Upper_Ontologies_Technical_Report_DISI-TR-06-21/links/0fcfd512d22286ae6de000000/A-Comparison-of-Upper-Ontologies-Technical-Report-DISI-TR-06-21.pdf (2007). Accessed 5 July 2021

Masolo, C. et al.: The {W}onder{W}eb Library of Foundational Ontologies.. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.11.4243 (2002). Accessed 5 July 2021

Monostori, L.: Cyber-physical production systems: roots, expectations and R&D challenges. In: Procedia CIRP, vol. 17, pp. 9–13. Elsevier, Amsterdam (2014). https://doi.org/10.1016/J.PROCIR.2014.03.115

Muller, H.A.: The rise of intelligent cyber-physical systems. Computer. **50**(12), 7–9 (2017). https://doi.org/10.1109/MC.2017.4451221

Ören, T. I., Ghassem-Aghaee, N., Yilmaz, L.: An ontology-based dictionary of understanding as a basis for software agents with understanding abilities. In Agent Directed Simulation Symposium, ADS 2007 – Proceedings of the 2007 Spring Simulation Multiconference, SpringSim 2007, pp. 19–27. Available at: https://www.researchgate.net/profile/Tuncer-Oeren/publication/322399131_An_Ontology-Based_Dictionary_of_Understanding_as_a_Basis_for_Software_Agents_with_Understanding_Abilities/links/5a577c10aca2726376b65703/An-Ontology-Based-Dictionary-of-Understanding- (2007). Accessed 7 July 2021

Pisch, F. Managing global production: theory and evidence from just-in-time supply chains Department of Economics, CEP Discussion Paper. Available at: https://www.alexandria.unisg.ch/260088/ (2020). Accessed 7 July 2021

Reinhart, G. et al.: Cyber-Physische Produktionssysteme – Produktivitäts- und Flexibilitätssteigerung durch die Vernetzung intelligenter Systeme in der Fabrik. *wt Werkstattstechnik online*, Jahrgang 1(H. 2):84–89 (2013)

Scheuermann, A., Hoxha, J.: Ontologies for intelligent provision of logistics services. ICIW 2012 : The Seventh International Conference on Internet and Web Applications and Services, 2(c), 106–111. Available at: http://72.52.166.99/articles/iciw_2012_4_20_20101.pdf (2012). Accessed 5 July 2021

Tantik, E., Anderl, R.: Potentials of the asset administration Shell of Industrie 4.0 for service-oriented business models. In: Procedia CIRP, pp. 363–368. Elsevier B.V., Amsterdam (2017). https://doi.org/10.1016/j.procir.2017.03.009

The LEVEL-UP Project.: Available at: http://www.levelup-project.eu/project (2019). (Accessed 28 July 2021)

Wache, H.: Semantische Mediation für heterogene Informationsquellen, KI. Akademische Verlagsgesellschaft Aka GmbH (DISKI Dissertationen zur Künstlichen Intelligenz), Berlin (2003)

Semantic Digital Twins for Retail Logistics

Michaela Kümpel, Christian A. Mueller, and Michael Beetz

Abstract As digitization advances, stationary retail is increasingly enabled to develop novel retail services aiming at enhancing efficiency of business processes ranging from in-store logistics to customer shopping experiences. In contrast to online stores, stationary retail digitization demands for an integration of various data like location information, product information, or semantic information in order to offer services such as customer shopping assistance, product placement recommendations, or robotic store assistance.

We introduce the semantic Digital Twin (semDT) as a semantically enhanced virtual representation of a retail store environment, connecting a symbolic knowledge base with a scene graph. The ontology-based symbolic knowledge base incorporates various interchangeable knowledge sources, allowing for complex reasoning tasks that enhance daily processes in retail business. The scene graph provides a realistic 3D model of the store, which is enhanced with semantic information about the store, its shelf layout, and contained products. Thereby, the semDT knowledge base can be reasoned about and visualized and simulated in applications from web to robot systems. The semDT is demonstrated in three use cases showcasing disparate platforms interacting with the semDT: Optimization of product replenishment; customer support using AR applications; retail store visualization, and simulation in a virtual environment.

1 Digitization of Stationary Retail

Participants of stationary retail, such as customers, store personnel, or store managers, demand for added service quality in their respective field of interest. While customers ask for shopping assistance or recommendations similar to online shopping experiences, store personnel wants to know the optimal time to restock

---

M. Kümpel (✉) · C. A. Mueller · M. Beetz
Institute for Artificial Intelligence, University of Bremen, Bremen, Germany
e-mail: michaela.kuempel@cs.uni-bremen.de; christian.mueller@cs.uni-bremen.de; beetz@cs.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_7

shelves with products and store managers want to determine the optimal product placement strategy for a store, for example.

As stationary retail is in the early stages of digitization, novel services are developed, aiming at satisfying the respective demands—mentioned above. Therein, models representing complex (retail) processes are introduced as a foundation for further analysis and reasoning. Based on a retail in-store logistics model (Kotzab and Teller 2005) and the Digital Twin model (Augustine 2020; Erkoyuncu et al. 2018; Grieves 2011), this chapter introduces a *semantic Digital Twin (semDT)* for retail as the semantic connection of a *scene graph* with a *symbolic knowledge base*, allowing for abstract reasoning about its contained facts in various retail applications. We refer to a *scene graph* (Costanzo et al. 2020; Mania and Beetz 2019) as a semantically annotated environment model that can be created automatically by robotic agents (Matsuo et al. 1999; Rusu et al. 2008; Sommer et al. 2019). The *symbolic knowledge base* consists of interlinked ontologies based on automated ontological models of everyday activities (Haidu and Beetz 2019) that contain further product or store logistic information using a uniform interface that supports integration of the scene graph.

The core entity in the semantic Digital Twin is the digital representation of a real store including its layout and offered products, augmented with various spatial, semantic as well as relational information. On the one hand, the scene graph holds information about a product location in relation to a shelf, a shelf layer, or other products on the same shelf layer and allows for optimization thereof. On the other hand, the symbolic knowledge base contains product information like 3D models of all objects, a product taxonomy, an ingredient classification as well as additional product information like product brand or awarded labels. Furthermore, the semDT knowledge base can be linked to store specific information like delivery or sales data. This knowledge can be reasoned about, visualized and modified in various environments from web interface to virtual environment, or applied on different platforms for *Augmented Reality* as well as robotic store assistance, for example. In *Augmented Reality* (AR), digital content is projected onto the real world (e.g., when using smart glasses like HoloLens) or a video stream of it (e.g., when using a Smartphone).

A robotic store assistant given the product replenishment task can access the semDT to combine location information, warehouse stock, as well as destination information. To further optimize the task, they can access semantic information and rules to determine the adequate time and sequence for restocking. Figure 1 illustrates the collaboration of robots using a semantic Digital Twin knowledge base. The robot to the left is creating the scene graph of a retail store by interlinking perceived object facts, thereby creating a semantic environment graph containing information about the product in relation to a shelf and shelf layer, for example: `<Product A> is_in <Shelf 3>, is_on <shelf_layer 2>`. The robot also captures empty *facings* and stores the information semantically (`<Product A> has_stock <0>`). We refer to a *facing* as the area between product separators on a shelf layer that can hold products subsequently behind each other. As all entities like shelf, shelf layer, and product also have positions relative to a reference frame,

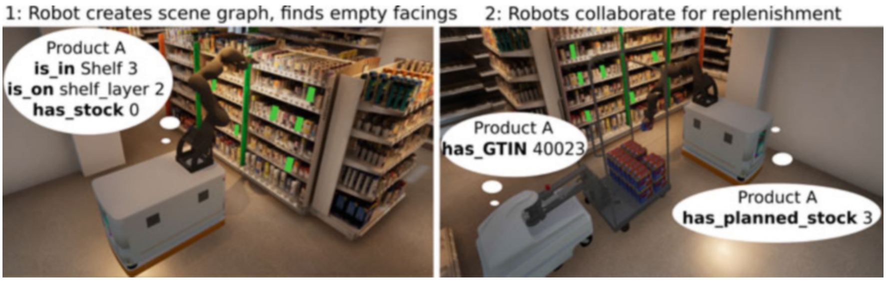


Fig. 1 Simulation of collaborating robot store assistants using a semantic Digital Twin

the environment knowledge can be shared between robotic agents as depicted in Fig. 1 to the right. From the inventory data in the scene graph, these two robots can infer which products need to be replenished (`<Product A>`), and from the symbolic knowledge base, they can derive how to identify them: `<Product A> has_GTIN <40023>`. The Global Trade Item Number (GTIN) is a unique identifier that is encoded in the product barcode. Since shelf layers are equipped with price tags containing this GTIN, they can be recognized by perception systems. From enterprise management systems, the semDT can include the planned stock for a given product (`<Product A> has_planned_stock <3>`, thereby having valuable information available to successfully collaborate for store assistance.

The connection of the scene graph to the symbolic knowledge base therefore makes it possible to reason about object properties in a shared environment, a benefit of implementing a semantic Digital Twin as knowledge base.

In this chapter, we define a semantic Digital Twin, discuss its composition and creation as well as applications of semDTs in retail stores to support users in their decision processes, and answer questions like “Which products need to be replenished in which quantity?,” “Which products are lactose-free?,” or “How does product placement influence sales data?.” Furthermore, we show the considerable economic potential of a semDT knowledge base in three example use cases:

Replenishment Process Based on inventory as well as delivery data and predefined rules, store personnel or robotic assistants can determine the adequate time, order, and destination for restocking of products.

Augmented Reality Shopping Assistant A customer can set product preferences on a HoloLens in order to highlight product information like an ingredient classification such as “vegan” or “lactose-free” to single out products that hold the given preference. Additionally, they can use their Smartphone to visualize interesting product information like awarded labels or contained allergens.

Digital Store Visualization and Robot Simulation A store manager can load a digital version of the store to visualize interesting information like deviations from planned to actual store layout. Aside from that, software developers demand for a simulation environment to test and verify their application before deployment to the real scenario.

The main contributions of the presented work are: (i) We introduce a semantic Digital Twin for retail as a connection of a scene graph and a symbolic knowledge base based on ontologies. (ii) We describe the semDT components and reasoning capabilities. (iii) We demonstrate the applicability of the semDT in three use cases for different users of the semDT.

2 Optimization of Retail Logistics

The advances in digitization have led to a tremendous growth of digital retailers and online shopping. Stationary retail needs to look for digital solutions in order to being able to compete with the information systems of their digital counterparts, providing an enhancement of the customer shopping experience and improving purchasing processes or optimizing the offered assortment depending on individual preferences and shopping behavior analysis. Contrary to digital retail, stationary retail lacks this transparency and abundance of information and thus has to focus on enhancing the customer shopping experience with more employees involved in customer interaction and care.

The introduction of digital technologies like Augmented Reality, Virtual Reality (VR), or mobile phone applications in retail stores can complement and enrich the shopping experience. This combination of digital technology in real environments can make shopping more appealing and entertaining for customers and accelerate localization or inspection for availability of searched products that contain certain ingredients or allergens, for example. This would allow for improved support for customers with allergies or other health conditions, goals, and needs. Digitization and the use of state-of-the-art technology can also support store personnel in their everyday work. Store employees can use AR devices to find product destinations for restocking or in order to complete *Click&Collect* orders. In *Click&Collect*, a customer can order products online to eventually collect it at their retail store in a preferred time slot. AR technology can also help train new employees faster by helping them in finding their way around the new store and searching for requested products.

These examples corroborate one problem that needs to be tackled for successful digitization of retail stores: Environment information plays a significant role in answering many questions in stationary retail, yet it is not included or linked to companies' knowledge bases. Therefore, in this chapter, we propose the semantic connection of environment information to product information in a semantic Digital Twin using an automatically created scene graph that is linked to a symbolic knowledge base consisting of interlinked ontologies.

Figure 2 visualizes example scenarios of this chapter. It illustrates the automatic scene graph creation by a robot, a Virtual Reality scene graph of a retail store, and an example application of the semantic Digital Twin: product placement investigation for assortment optimization.

Fig. 2 Example scenarios of the semantic Digital Twin

The semantic Digital Twin enables retail users, e.g., store managers, to visualize all products from a certain brand. They can ask for the sales quantity of these products and compare different product placements in regard to product sales data. Since stores are linked to retail chain and company-specific product information, one can even compare different store layouts to find similarities and differences.

Semantic Digital Twins can contribute to solve various problems, particularly in logistic processes:

Store managers are interested in questions such as “Which products need to be replenished in which quantity?” or “Which products are sold out?” at any time.

Retail store personnel or robotic store assistants can pose questions such as “Where should this product be sorted?” or “At what time should which product be restocked?.”

Customers may want to know “Where can I find the products from my shopping list?” or “Which products are produced economically friendly?.”

Regional managers might be interested in questions such as “What experiences do stores have with the placement of certain products in different areas?” or “Which store sells which products?.”

Logistics managers can get answers to questions like “How much space is available in the store for a certain product?” or “In which order should products be palletized to make unloading and refilling of the shelves as efficient as possible?.”

Marketing specialists can use semDTs to compare and evaluate store layouts and assess how product positioning influences sales figures.

Supply chain managers get optimized stock information to abbreviate order and delivery times and minimize warehouse stock.

Software developers in the field of robotics can use semDTs to simulate robot behavior before applying it to the real robot in the store.

These are only a few of the applications the semantic Digital Twins can contribute to. As development and experience with this system move on, further applications and use cases will arise.

3 Semantic Digital Twin: A Digital Representation of a Retail Store

The concept of a

> *Digital Twin*, the digital equivalent to a physical product,

was introduced in 2003 (Grieves 2011) and later defined variably in related work with different foci in a manufacturing context like applications in NASA and US Air force vehicles (Glaessgen and Stargel 2012) or as a living model of a physical asset (Liu et al. 2018). If we focus on the retail sector instead of manufacturing, the object of interest changes from physical item or assembly to an ecosystem centered around purchasable products such as described in in-store logistic models (Kotzab and Teller 2005; Zheng et al. 2019). We propose embedding of a Virtual Reality scene graph representation of the physical elements, allowing for integration of physics simulation in addition to visualization. We follow the approach of Erkoyuncu et al. (2018) to

> connect Digital Twin and physical asset through their interrelations,

thereby highlighting the importance of relations and semantics. Further inclusion of non-physical entities and their properties enables complex reasoning capabilities. Hence, we define a semantic Digital Twin as:

Definition 1 (Semantic Digital Twin: semDT) A semantic Digital Twin is a symbolic representation of robots, human beings, and their environment as physical elements connected to complementary non-physical entities as well as their properties and interrelations, represented by data structures of Virtual Reality scene graphs. Thereby abstract information associated with the entity of interest can be inferred, reasoned about, and visualized through a variety of media to predict current or future conditions. Particularly, actions can be simulated, and hypothetical scenes can be rendered to support and enhance decision-making.

We can use semantic Digital Twins to store, interpret, and query information of heterogeneous resources from enterprise data over web information to simulated processes or effects of actions through a uniform interface as depicted in Fig. 3. Its symbolic knowledge base is based on an information system for storage, management, and usage for embodied intelligent agents (Beßler et al. 2020) and,

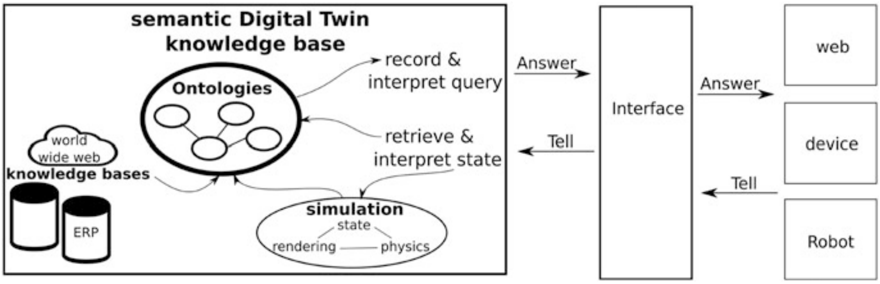


Fig. 3 Semantic Digital Twin model and its main components

therefore, describes contained facts about products and logistics processes in a machine-understandable way, enabling many new applications in the first place. The components of the semDT will be further explained in Sect. 4.

The semDT realistically represents the physical store, including 3D images for visualization and semantic information. Each element is uniquely identified so that it can be linked to additional knowledge bases.

Semantic Digital Twins can be automatically created by perception-based systems (Arpenti et al. 2020). Inclusion of merchandise management systems (ERP) can allocate sales information and automatically connect it to object location information, enabling system users to answer queries for product sales per shelf, brand, or store. The semDT combines disjointed business data into a coherent picture. The entire company thus has access to the most up-to-date and complete data stock. In this way, logistics can be individually coordinated for each store and the flow of goods can be sorted on site according to an optimized placement of products. Simulation, on the other hand, facilitates extensive marketing studies and supports the integration and control of various systems such as service robots through the included semantic information.

Subsequently, semDTs can function as information carriers with a unique level of detail, visual realism, and cognitive capabilities for the development of innovative information services.

4 Building Blocks of the Semantic Digital Twin

The semantic Digital Twin is based on environment information perceived by a robot. This environment information is transformed into a scene graph, a semantically annotated environment model. Thereby a product is assigned a position relative to other products on a shelf and shelf layer, allowing reasoning about its location. For creation of the semDT, the scene graph is connected to a symbolic knowledge base containing enterprise product information like product brand or price but also product information like an ingredients classification or label information from various sources as depicted in Fig. 3, enabling the semDT to quickly adapt to changing product information or consumer needs.

The semantic Digital Twin can be described as a Knowledge Representation and Reasoning (KRR) framework, which is a knowledge base that organizes information of different knowledge sources from the literature, perception system, historical data as well as forecasts and other open research data services. In the following, we continue to describe the knowledge representation of the semantic Digital Twin and the creation of its components, the scene graph, and the symbolic knowledge base. The section closes with an overview of the reasoning capabilities of the semantic Digital Twin, which enable answering of the questions posed in Sect. 2.

4.1 Knowledge Representation

The semantic Digital Twin represents its knowledge in the form of ontologies; therefore, all information is stored using triples of entities and their relations. Ontologies such as information sources provide meaning to the contained facts and entities and allow reasoning about them (Staab and Studer 2010). Each ontology comprises information from a different source or with a different focus in order to make them modular and interchangeable. As depicted in Fig. 4, the product is the central object in each ontology and therefore used for interlinking between the ontologies. It can easily be expanded by any ontology that uses a GTIN as product intendifier. In this example, five ontologies are shown, containing location and inventory information (`<Product A> is_in <Shelf 3>, has_stock <2>,...`) as the scene graph that is connected to the symbolic knowledge base with enterprise information (e.g., `<Product A> brand <Somat>`) as well as additional product information derived from different sources from robot to web information (`<Product A> is_a <detergent>, has_label <vegan>, has_ingredient <perfume>`). This information can be used as a basis for many consumer applications.

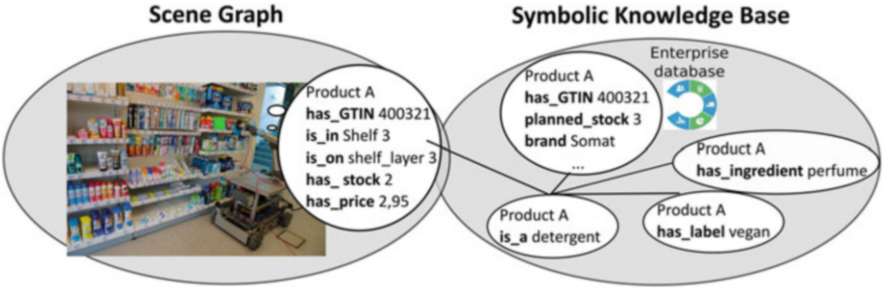


Fig. 4 Example of linked knowledge sources in the semantic Digital Twin

In order to achieve a rich foundation for various reasoning tasks, the semDT KRR framework maintains a virtual replica of the retail store, which is enriched by ontologies comprised of expert knowledge, current state information captured from mapping and sensing technologies as well as interesting product information. Given such foundation, the framework allows investigating and mining hypotheses to support queries of retail business actors and Big-Data-enhanced abstract reasoning tools; for instance, retail-relevant concepts, such as “misplaced objects,” can be inferred if differing products are detected in a facing. Furthermore, responses and current state information are intuitively rendered to aid and assist in decision-makings for everyday tasks in a retail store.

KnowRob: Knowledge Processing Framework
The ontology-based KnowRob system, first introduced in 2009 (Tenorth and Beetz 2009) and extended in 2018 (Beetz et al. 2018), is at the forefront of cognitive robot control in the household domain in terms of the extent of information its knowledge base represents (Thosar et al. 2018). It combines encyclopedic knowledge with implicit knowledge but also techniques for acquiring knowledge and for grounding it in a physical system. KnowRob can serve as a common semantic framework for integrating information from different sources. Thereby, it combines robot sensor input with static encyclopedic knowledge, common-sense semantic knowledge, task descriptions, virtual environment models, and object information. Furthermore, it supports different deterministic and probabilistic reasoning mechanisms, clustering, classification and segmentation methods and includes query interfaces as well as visualization tools. As a result, we propose KnowRob as a suitable knowledge processing framework for the semantic Digital Twin.

4.2 Scene Graph

The KnowRob ontologies used for scene graph creation are built on top of existing ontologies. They include concepts about objects from the Everyday Activity Science and Engineering (EASE) framework (Bateman et al. 2017). These concepts are connected to a general description of actions, tasks, and agents from the Socio-physical Model of Activities (SOMA) (Beßler et al. 2020). Lastly, concepts about participants in events or localization of objects are integrated from the DOLCE+DnS Ultralite ontology (Gangemi et al. 2002). Thus, the KnowRob ontologies link objects to tasks of actions with participants as a basis for interpretation and reasoning. The KnowRob ontologies are expanded for the semDT by retail-relevant ontologies containing store or layout information with descriptions about the existing shelf systems, their features (height, width, the number of shelf layers) as well as object positions and product information (barcode, stock). All objects in the store are represented as concepts with properties, defining how they are related to each other. For example, shelf layers support a certain weight, leading to a maximum number of products that can be placed on it.

Retail stores are structured environments containing systematically organized shelf systems. The structural design of these organized shelf systems can be formalized using rules in predicate logic (Rashmi and Rangarajan 2018), enabling reasoning over the contained entities but also automatic creation of semantic environment maps by robots. Such a semantic environment map contains a rich description of objects in the stationary environment and can be used to answer queries like "Where are deodorants located?" or "How does shelf X differ from its abstract representation in the planned store layout?." Due to the fact that the semantic environment map links to concepts about actions, objects, and even actions effects, we refer to it as a scene graph.

Internally, KnowRob generates a hierarchical structure of the asserted objects, while a robot is scanning the store for scene graph creation. Together with the semantic knowledge about these objects and their 6D pose, KnowRob creates a virtual snapshot of the world's state as depicted in Fig. 2.

Automatic Scene Graph Creation by Robots

We apply robots to automatically create the scene graph for the semantic Digital Twin, but the scene graph can be created by any perception-based system. The process of scene graph creation consists of two major steps: *layout detection* and *store monitoring*. During layout detection, rarely changing features of the store (like room size and shelf positions) are captured. This task needs to be performed once for each new store layout. The store monitoring process is the repeating process of stocktaking.

Layout Detection The layout detection starts with the creation of a 2D map of the store using grid mapping as simultaneous localization and mapping technique (Grisetti et al. 2007). Afterward, the robot drives through the store to create a basic scene graph of the environment without product information. Each shelf is detected using a Quick Response (QR) code. The position data is added to the scene graph in such a way that shelf positions in relation to other shelves or points of interest can be calculated and reasoned about.

Store Monitoring During store monitoring, frequently changing product positions are detected automatically by the robot. For each shelf, the robot scans the shelf vertically to detect shelf layers and horizontally for each shelf layer to detect price labels and product separators. The stock as the number of products between two product separators is estimated based on detected product features of an RGB-D camera similar to other approaches (Donahue et al. 2014). Each price label contains a barcode that encodes the GTIN of a product, thereby adding product information to the scene graph. This product information is continuously updated based on sales data. If the store monitoring process is performed regularly, the scene graph of two different time points can be compared to detect irregularities between calculated inventory and actual inventory that can be reasoned about.

4.3 Symbolic Knowledge Base

If a robotic agent is given the task to pick objects in a retail store in order to fulfill a customer shopping order and an unexpected situation is experienced, such as a misplaced product (e.g., a “coke light” bottle in the facing of “coke” bottle), the robot needs to be able to reason about the situation using externally acquired knowledge (e.g., current store layout, product classification) in order to successfully finish the task.

As previously mentioned in Sect. 4.2, the scene graph is created using the ontology-based KnowRob system. The ontologies in the symbolic knowledge base can optionally be integrated in the KnowRob system. KnowRob can simplify the process of integrating information from different sources and translating them to ontologies. This is specifically useful for the semantic Digital Twin since most retail companies store their data in merchandise management systems (ERP). With KnowRob, product classification information from an ERP system can be translated into a product taxonomy in Web Ontology Language (OWL) format, for example. Furthermore, KnowRob is able to integrate *Semantic Web* information.

The Semantic Web (Berners-Lee et al. 2001) is an extension of the World Wide Web (Sirin et al. 2003), aiming at structuring content of web pages by using standardized, machine-readable formats to represent entities, their properties, and relations so that it can be interpreted by software agents or robots. This integration enables the semantic Digital Twin to compete with online stores and their recommender systems. The label and ingredient information depicted in Fig. 4 are based on Semantic Web information, for example. The ingredients ontology can further be connected to an allergen ontology, enabling the semDT to reason about contained allergens in the products available at the store as shown in an example query asking for all products that contain an ingredient that is classified as being hazardous to the environment in Fig. 5.

```
?- sparql_query(’
PREFIX tax: <http://knowrob.org/kg/ProductTaxonomy.owl#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX in: <http://knowrob.org/kg/Ingredients.owl#>
PREFIX all: <http://knowrob.org/kg/allergen.owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?allergen ?product ?class WHERE {
?allergen all:has_depiction all:hazardous_to_the_environment.
 ?ingredient owl:sameAs ?allergen.
?product in:has_ingredient ?ingredient.
?product owl:sameAs ?prod.
?prod rdf:type ?class.
 FILTER (?class != owl:NamedIndividual && ?class!= tax:Product)}’,
 Row, [endpoint(’URL of SPARQL endpoint’)]).
```

Fig. 5 Example Prolog query for all products and their classification that have an ingredient that is classified as hazardous to the environment

Table 1 Results excerpt (out of 119 results[a]) for the query in Fig. 5

depiction	env. hazard allergen	contained by product	product classification
	all:Benzyl_benzoate	in:3614224040116	tax:deodorant
	all:Benzyl_benzoate	in:4001638086561	tax:facial_care
	all:Benzyl_benzoate	in:4005900471109	tax:bodylotion
	all:Benzyl_cinnamate	in:8712561521840	tax:deodorant
	all:Farnesol	in:4001638080262	tax:facial_care
	all:Farnesol	in:7317400017303	tax:tooth_care
	all:Hexyl_cinnamal	in:3600540901187	tax:deodorant
	all:Methylisothiazolinone	in:4015000960847	tax:detergent
	all:Natriumhypochlorid	in:4032651214181	tax:first_aid
	all:Petrolatum	in:3574661199634	tax:baby_care

[a] All queries presented in this work are accessible online: `http://grlc.io/api/K4R-IAI/NonFoodKG/SPARQLfiles`

The depicted *Prolog* query is using the Semantic Web library package; *Prolog* is a logic programming language (Nilsson and Małuszyński 1990) used for reasoning in the symbolic knowledge base. In this figure, a *SPARQL* query at a specific SPARQL endpoint is called (lines 1 and 13). *SPARQL* is a graph-based query language and recommended query language for Resource Description Framework (RDF) graphs (Angles and Gutierrez 2008; Pérez et al. 2006). Lines 2 to 6 declare prefixes to be used in the SPARQL query given in lines 7 to 12. The query asks for all products that contain an ingredient that is classified as being hazardous to the environment, a label that has the standard depiction shown in Table 1 to the left, which is referenced in Semantic Web sources like Wikidata, for example. Line 7 names the variables that are to be returned by the query: `?allergen`, the ingredient contained by a product and classified as hazardous to the environment, `?product`, the GTIN identifier of the product, and `?class`, the class of the product. Line 8 queries the allergen ontology for all ingredients that have the relation `<?allergen> has_depiction <hazardous_to_the_environment>`. Line 9 links the allergen to the ingredients ontology using the `owl:sameAs` statement. Then all products that contain any of the returned ingredients are searched for in line 10 (`<product> has_ingredient <ingredient>`). In line 11, the resulting product list is linked to the product taxonomy through the use of `owl:sameAs` statements. Lastly, the product class is retrieved in line 12 (`<?prod> type <?class>`). Line 13 filters the result set for redundant classes.

An excerpt of the results for the query in Fig. 5 is shown in Table 1. It depicts the international pictogram for environmental hazard in column 1. The classification of ingredients as being hazardous to the environment was extracted from the European chemicals agency ECHA website[1] and integrated into the semDT allergen classification. All ingredients that are listed as environmental hazard allergen in the table, column 2, therefore are classified according to the ECHA website. Column 3 contains the GTINs of the products that contain the hazardous ingredient, and column 4 contains a product classification. The prefixes in columns 2 to 4 show

---

[1] European chemicals agency website: https://www.echa.europa.eu/.

that all information is extracted from different sources, emphasizing the benefits of reasoning over connected ontologies in the semDT. Using Prolog and ontologies, information sources can easily be included or exchanged and reasoned about.

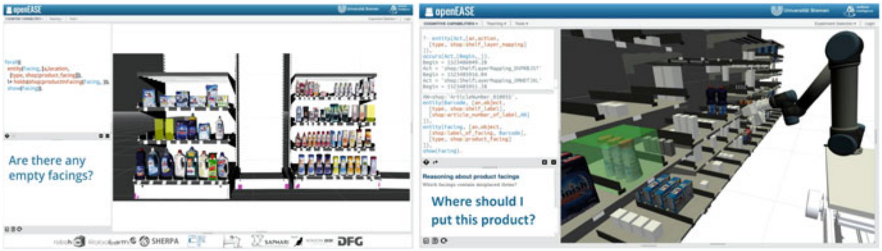
Consumer applications based on this symbolic knowledge base are further elaborated in the use case described in Sect. 5. If customers particularly prefer environmentally friendly produced products, they can use the symbolic knowledge base to filter out products with non-compliant or hazardous ingredients as shown in the example above, for instance.

Entity Linking in the Semantic Digital Twin
We state that ontologies can easily be included or exchanged in the semDT. This is true for all ontologies that hold product information and include a GTIN as unique identifier for the contained products, which is the case for most ontologies in the semDT. Entity linking in knowledge bases is a widely studied research problem (Shen et al. 2014), even more so for web-scale data (Lin et al. 2012). For all semDT ontologies, we use string matching techniques to identify entities that are to be linked. `"Benzyl_benzoate"` from the allergen ontology thereby is linked to `"benzyl_benzoate"` and `"benzylbenzoat"` from the ingredients ontology, for example.

4.4 *Reasoning*

Reasoning about semantic information in conjunction with environment information is the main capability of the semantic Digital Twin.

As previously mentioned, the logic programming language *Prolog* is used to reason about information contained in the symbolic knowledge base (see Sect. 4.3). With Prolog, we are able to answer questions as posed in Sect. 2 and develop applications that are based on it like in the use cases presented in Sect. 5. Since the semDT also stores 3D models of the perceived objects and connects scene graph with symbolic knowledge base, we can generate a web user interface that highlights information on demand as depicted in Fig. 6.



Are there any empty facings?

Where should I put this product?

Fig. 6 Example reasoning capabilities of the semantic Digital Twin

```
1    ?- GTIN=shop:'ArticleNumber_GTIN',      #input GTIN
2    entity(Barcode,[an,object,              #retrieve object named Barcode
3    [type, shop:shelf_label],               #Barcode is a shelf_label
4    [shop:article_number_of_label,GTIN],    #Barcode has indentifier GTIN
5    ]),
6    entity(Facing,[an,object,               #retrieve object named Facing
7    [shop:label_of_facing, Barcode],        #Facing has_label Barcode
8    [type, shop:product_facing],            #Facing is a product_facing
9    ]),
10   show(Facing).                           #highlight Facing
```

Fig. 7 Example Prolog query for destination of a given product GTIN

Figure 6 displays the web interface openEASE,[2] an open access web-based knowledge service based on KnowRob, where users can upload episodes of data and query it for information. OpenEASE facilitates access to the knowledge base using Prolog. The Prolog query asking for empty facings is depicted in Fig. 6 to the left, and the Prolog query asking for a specific product destination is visualized in Fig. 6 to the right and generalized with comments in the following Fig. 7.

The query in Fig. 7 requires an input parameter specified in line 1, which is the GTIN of a product. The GTIN is used to first retrieve the corresponding `Barcode` object from the `shop` ontology in lines 2 to 5. The `Barcode` variable is defined as an `<object> type <shelf_label>` in line 3 with the GTIN uniquely identifying the object as of line 4 (`article_number_of_product <GTIN>`). The `Barcode` is used as input to retrieve the destination facing ob the object in lines 6 to 9. The variable `Facing` is defined as a `product_facing` in line 8, and the facing corresponding to the input product GTIN saved in the variable `Barcode` is scanned for in line 7: `<Facing> label_of_facing <Barcode>`. Lastly, the facing is being highlighted in line 10 using the `show` command.

OpenEASE offers a variety of visualizations of the knowledge base content, including 3D visualizations of the store including its entities (e.g., shelf, products, facings, or robots as in Fig. 6), the trajectories they perform as well as interval logic layers for showing temporal relationships between different events. In order to exchange information between the semDT and users, openEASE serves as an intuitive interactive interface providing information about the current state of the store and facilitating Q&A queries. Responses can be rendered in the form of physical or abstract entities to assist humans as well as robots in everyday retail tasks as will be demonstrated in the use case section.

---

[2] OpenEASE web interface: data.open-ease.org/.

Reasoning About Action Information

A main advantage of the ontologies used in KnowRob is their integration of activity knowledge from the SOMA ontology (Beßler et al. 2020). In combination with the scene graph based on object information from the EASE framework (Bateman et al. 2017) and task descriptions from the DOLCE+DnS Ultralite ontology (Gangemi et al. 2002), KnowRob can reason about performed actions by different agents. This reasoning capability can also be accessed via the openEASE web interface.

Action information in KnowRob is stored as episodic memories (Bartels et al. 2019) based on general plan descriptions (Koralewski et al. 2019). These general plan descriptions follow an action and task hierarchy where a stocktaking action performed by a robot can be comprised of different tasks like driving, positioning, and scanning. Figure 8 shows a general plan description for the action of looking at a product, a subaction of the stocktaking action. Figure 9 describes the plan for a navigation action to a destination.

Additionally to the action hierarchy, episodic memories can store agent information as participant or performer of an action, assign roles to objects for a certain action (e.g., `shelf 2` in the general plan in Fig. 9 would be assigned the role *destination*), and log times and duration of the performed actions.

Figure 10 shows example action information about a recorded stocktaking episode by a robot. It shows that the stocktaking action is divided into various stocktaking tasks for the different shelves and shelf layers, respectively. All scanned facings are highlighted in green, and empty facings in yellow. Additionally, robot movement is logged as “AssumingArmPose” and “LimbMotion.” One can inspect each task and action for their duration as well as participants of the action or action effects. The event can even be replayed in an interactive simulation.

Action information cannot only be recorded for robots, the semDT can also store anonymous episodic memory information about shopping events using AR devices or sensors in the retail stores. Such information is a valuable resource that can be applied to various use cases such as to optimize robotic store assistants as well as customer experience.

Fig. 8 Plan description for LookingAt action

```
(an action
  (type LookingAt)
  (an object
     (type product)))
```

Fig. 9 Plan description for Navigation action

```
(an action
  (type Navigation)
  (to (a location
     (in front of (an object
          (name shelf 2))))))
```

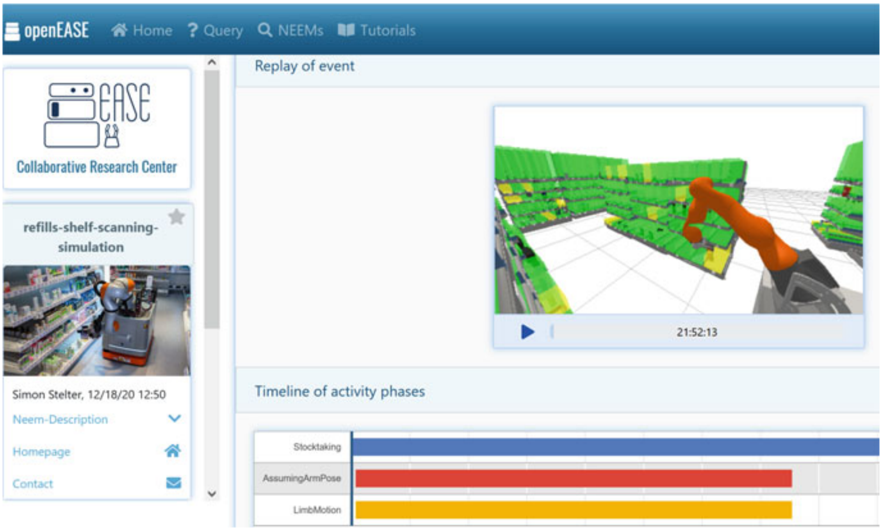


Fig. 10 Example action information in the semantic Digital Twin

5 Semantic Digital Twin Use Cases in Retail Logistics

We demonstrate the applicability of a semantic Digital Twin in three use cases based on different users of the system. The semDT provides vast retail-relevant information through machine-understandable data formats and interfaces. This enables services including AR applications and robots to access information, such as poses of products and shelves or properties of objects like product dimensions, in order to perform tasks such as shelf replenishment.

As discussed in Sect. 2, store personnel and robotic store assistants aim at optimizing logistics such as for product replenishment that is demonstrated in use case 1. Use case 2 shows customer AR applications accessing the semDT for product-specific information like contained ingredients or awarded labels. Lastly, use case 3 demonstrates how a digital store can be visualized in a virtual environment, where store managers can simulate action effects.

5.1 Use Case 1: Replenishment Process

This use case demonstrates the benefits of a semantic Digital Twin from the perspective of store personnel and robotics using the example task of product replenishment.

5.1.1 AR Supported Replenishment

Product replenishment poses several challenges: It is required to be performed accurately and regularly in each retail store. New employees need to learn product destinations before they are able to efficiently perform replenishment, but all store personnel need time to adapt to changing store layouts, seasonal product stands, and new offerings.

The semDT can be accessed by AR applications for store personnel to ease the task of replenishment and order fulfillment, for example. Figure 11 depicts such an AR app where store personnel can ask for the location of a product group as in "Where is dishwashing detergent?." The app locates the AR device in its environment using persistent digital environment anchors in the store. Then the camera view of the environment is augmented with digital waypoints that are aligned relative to these anchors, thereby leading the user to the product destination. The correct path is recalculated regularly to account for direction changes of the device. In order for app users to distinguish waypoints from destination points, they are assigned different colors. While green waypoints mark the path, an orange waypoint marks the searched product destination, see in Fig. 11 the shelf holding dishwashing detergent products.

Fig. 11 AR routing for store personnel

Fig. 12 A robot identifying empty facings in simulation (left) and replenishing products in the store accordingly (right)

5.1.2 Robotic Replenishment

Robotic store assistants can also be used to contribute to product replenishment. They require reliable environment information for this task. The main benefit of the semantic Digital Twin is its connection between environment information in the form of a scene graph and semantic information in its symbolic knowledge base. As a proof of concept, this use case shows the replenishment process performed by a robot based on scene graph information as depicted in Figs. 2 and 12. The robot responsible for replenishment can access the scene graph information and query the semDT for all empty facings as shown in Fig. 6 to the left. By accessing the symbolic knowledge base of the semDT, the robot can further reason about the products (e.g., their weight and size and how these properties affect the replenishment task) and the quantity of products that need to be fetched from the warehouse. Afterward, they can query the semDT for the respective product destinations as described in Fig. 7. Subsequently, the robot can be ordered to move to an item, localize the item using its camera, and collect the item with its gripper for replenishment purposes. Figure 12 shows a robot querying the semDT for empty facings in the simulation interface and accordingly performing the replenishment in a real store scenario.

5.2 Use Case 2: Augmented Reality Shopping Assistant

This use case demonstrates the benefits of a semantic Digital Twin from a customer perspective in two AR applications on different devices: a HoloLens and a mobile phone.

5.2.1 HoloLens Application

Customers demand information and recommendations as presented in online shopping experiences. The semDT supports an AR app designed for the AR glasses HoloLens, highlighting individual preferences in a retail store. An example query

Fig. 13 AR apps to highlight consumer preferences like ingredients, label, or hazard information

that could be used in this AR application for products that contain ingredients that have a classification of being hazardous to the environment has been discussed in Sect. 4.3. In the same manner, AR applications can access the semDT and query for all products that contain preservatives, as shown in Fig. 13 to the left. All products of the store are aligned relative to shelves using scene graph information and environment anchors as in use case 1. If a customer checks a preference in the menu, all products that contain the ingredient are overlayed by a red "X," indicating that the product behind the overlay contains the ingredient and is not intended to be bought. This allows for customers to inspect all remaining products.

5.2.2 Mobile Phone Application

The semDT can just as well be accessed by other devices. We present a second AR app running on a mobile phone, using object recognition to identify products in the store. If customers are interested in a product, they can point the camera toward the interesting product. Once the product is identified, interesting product information like awarded label and hazard information is being displayed, as depicted in Fig. 13 to the middle and right. If consumers are interested in further information, they can swipe the product name that appears at the bottom of the screen and a second screen with additional product information is being displayed.

5.3 *Use Case 3: Digital Store Visualization and Robot Simulation*

This use case demonstrates the benefits of a semantic Digital Twin from the perspective of a store manager to visualize store information and software developers to simulate robot behavior.

5.3.1 Semantic Digital Twin Visualization

SemDT users like store managers, who want to investigate available product information in order to, e.g., optimize product placement, face the challenge to virtually imagine the appearance of the store. They require a mechanism to spatially represent and visualize different states and configurations of the store layout in order to efficiently optimize logistic store processes such as product placement. Therefore, in addition to the web-based interface openEASE, the semDT provides a photo-realistic 3D visualization of the semDT as an extension of the Virtual Reality scene graph, facilitating users to interact with the semDT in an intuitive way. A store manager can visually inspect, compare, and analyze different store layout configurations or move articles and change layouts in different states of the semDT. Using Prolog queries, travel paths and trajectories can be visualized. In Fig. 2, we replicated the example store located at the University of Bremen. Figure 14 demonstrates a visualization of empty facings in the virtual semDT.

Using the Prolog query interface, semDT users can also highlight interesting information as depicted in Fig. 15, where all products containing hazardous ingredients that are reachable by children are highlighted. This query uses the query for hazardous ingredients as shown in Fig. 5 and filters out all products that are below a given height and therefore reachable by children.

Fig. 14 Visualizing empty facings in the virtual semDT

Fig. 15 Highlighting "hazardous products reachable by children" in the virtual semDT

5.3.2 Robot Simulation

Simulation is crucial for the development of novel retail applications. For instance, software developers of robot applications might have no or only limited access to retail stores. Furthermore, robotic applications tend to be complex systems, consisting of perception, localization, and control frameworks, among others. By offering interfaces equivalent to the real robot interfaces, it is possible to execute and test the application in the virtual semDT before deploying it in the retail store. Examples of different robot simulations are shown in Figs. 1 and 16. Figure 16

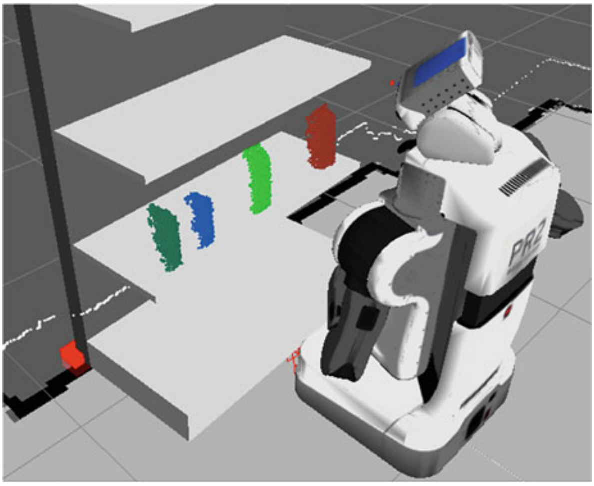


Fig. 16 Robot simulation with semDT

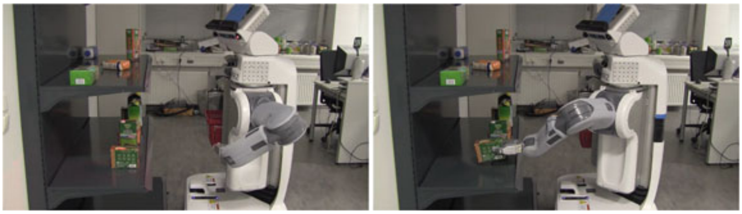


Fig. 17 Real robot perceiving and picking objects based on simulation results

simulates a robot perceiving objects on a shelf layer, whereas Fig. 17 shows the real robot perceiving and picking the same objects in a retail store.

6 Conclusion

In this chapter, we introduced the semantic Digital Twin for Retail Logistics as a semantically enhanced digital representation of a retail store and the necessary connection between environment information in a scene graph and semantic product information in a symbolic knowledge base based on ontologies that allow for visualization, simulation, and complex reasoning tasks. This chapter outlines the key components and potential benefits of a semDT. With the semDT, store personnel can derive product destinations and implement rules for inferring the optimal order of product replenishment. Customers can use the semDT AR applications on various devices like HoloLens or mobile phone for an enhanced shopping experience, perceiving vast product information as well as visual and tactile sensation. Store managers can access the semDT for layout inspection or optimization, and software developers in the field of robotics can simulate robot behavior in the semDT.

Another contribution of this chapter is a demonstration of the capabilities of the semDT. The semDT ontologies derived from various sources allow for complex reasoning, which was highlighted in example queries from different platforms and applications. We demonstrated the applicability of the semantic Digital Twin in three use cases focusing on different users of the system (i.e., store personnel, manager, customer as well as software developer) accessing the semDT knowledge via various devices.

7 Future Directions

The proposed semantic Digital Twin aims at providing mechanisms offering further transparency for logistic retail processes of a store; further development and extensions of semDT capabilities are sought. Besides implementation of semDTs for different retail chains, like drugstores, grocery stores, and bookstores, which are able to afford high investments in digital infrastructure, we aim at developing semDTs in diverse and small retail business like rural corner shops as well. Since the semDT ontologies are created in a modular fashion, they can be used in different applications and can be integrated into online shops, which will be implemented in a future version of the semDT. Another future direction is set on the interconnectivity of semDTs enabling new possibilities of such knowledge transfer among different retail domains to optimize logistic processes, for example.

Acknowledgments We thank Kaviya Dhanabalachandran, Alina Hawkin, Nils Leusmann, Michael Neumann, and Hoang Giang Nguyen for their contribution to this work.

The research reported in this paper has been partially supported by the Federal Ministry for Economic Affairs and Energy BMWi within the Knowledge4Retail project, subproject semantic Digital Twin 01MK20001M (https://knowledge4retail.org) as well as the German Research Foundation DFG, as part of Collaborative Research Center (Sonderforschungsbereich) 1320 "EASE—Everyday Activity Science and Engineering," University of Bremen (http://www.ease-crc.org/). The research was conducted in subproject R03.

References

Angles, R., Gutierrez, C.: The expressive power of SPARQL. In: International Semantic Web Conference, pp. 114–129. Springer, Berlin (2008)

Arpenti, P., Caccavale, R., Paduano, G., Fontanelli, G.A., Lippiello, V., Villani, L., Siciliano, B.: RGB-D recognition and localization of cases for robotic depalletizing in supermarkets. IEEE Robot. Automat. Lett. **5**(4), 6233–6238 (2020)

Augustine, P.: Chapter four – the industry use cases for the digital twin idea. In: Raj, P., Evangeline, P. (eds.) The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases. Advances in Computers, vol. 117, pp. 79–105. Elsevier, Amsterdam (2020). https://doi.org/10.1016/bs.adcom.2019.10.008. http://www.sciencedirect.com/science/article/pii/S0065245819300610

Bartels, G., Beßler, D., Beetz, M.: Episodic memories for safety-aware robots. KI-Künstliche Intelligenz **33**(2), 123–130 (2019)

Bateman, J., Beetz, M., Beßler, D., Bozcuoğlu, A.K., Pomarlan, M.: Heterogeneous ontologies and hybrid reasoning for service robotics: the EASE framework. In: Iberian Robotics Conference, pp. 417–428. Springer, Berlin (2017)

Beetz, M., Beßler, D., Haidu, A., Pomarlan, M., Bozcuoğlu, A.K., Bartels, G.: Know Rob 2.0—a 2nd generation knowledge processing framework for cognition-enabled robotic agents. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 512–519. IEEE, Piscataway (2018)

Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Sci. Am. **284**(5), 34–43 (2001)

Beßler, D., Bozcuoğlu, A.K., Beetz, M.: Information system for storage, management, and usage for embodied intelligent systems. In: Information Storage, pp. 135–159. Springer, Berlin (2020)

Beßler, D., Porzel, R., Pomarlan, M., Vyas, A., Höffner, S., Beetz, M., Malaka, R., Bateman, J.: Foundations of the socio-physical model of activities (SOMA) for autonomous robotic agents. arXiv preprint arXiv:2011.11972 (2020)

Costanzo, M., Stelter, S., Natale, C., Pirozzi, S., Bartels, G., Maldonado, A., Beetz, M.: Manipulation planning and control for shelf replenishment. IEEE Robot. Automat. Lett. **5**(2), 1595–1601 (2020)

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: DeCAF: a deep convolutional activation feature for generic visual recognition. In: ICML'14: Proceedings of the 31st International Conference on International Conference on Machine Learning, pp. 647–655. PMLR (2014)

Erkoyuncu, J.A., Butala, P., Roy, R., et al.: Digital twins: understanding the added value of integrated models for through-life engineering services. Proc. Manuf. **16**, 139–146 (2018)

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening ontologies with DOLCE. In: International Conference on Knowledge Engineering and Knowledge Management, pp. 166–181. Springer, Berlin (2002)

Glaessgen, E., Stargel, D.: The digital twin paradigm for future NASA and US Air Force vehicles. In: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, p. 1818 (2012)

Grieves, M.: Virtually Perfect: Driving Innovative and Lean Products Through Product Lifecycle Management. Space Coast Press, Brevard (2011)

Grisetti, G., Stachniss, C., Burgard, W.: Improved techniques for grid mapping with Rao-Blackwellized particle filters. IEEE Trans. Robot. **23**(1), 34–46 (2007)

Haidu, A., Beetz, M.: Automated models of human everyday activity based on game and virtual reality technology. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 2606–2612. IEEE, Piscataway (2019)

Koralewski, S., Kazhoyan, G., Beetz, M.: Self-specialization of general robot plans based on experience. IEEE Robot. Automat. Lett. **4**(4), 3766–3773 (2019)

Kotzab, H., Teller, C.: Development and empirical test of a grocery retail instore logistics model. Br. Food J. **107**(8), 594–605 (2005)

Lin, T., Etzioni, O., et al.: Entity linking at web scale. In: Proceedings of the Joint Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction (AKBC-WEKEX), pp. 84–88 (2012)

Liu, Z., Meyendorf, N., Mrad, N.: The role of data fusion in predictive maintenance using digital twin. In: AIP Conference Proceedings, vol. 1949, p. 020023. AIP Publishing LLC, Melville (2018)

Mania, P., Beetz, M.: A framework for self-training perceptual agents in simulated photorealistic environments. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 4396–4402. IEEE, Piscataway (2019)

Matsuo, T., Tanaka, K., Abe, N.: Automatic map generation with mobile robot. In: IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028), vol. 4, pp. 680–685. IEEE, Piscataway (1999)

Nilsson, U., Małuszyński, J.: Logic, Programming and Prolog. Wiley, Chichester (1990)

Pérez, J., Arenas, M., Gutierrez, C.: Semantics and complexity of SPARQL. In: International Semantic Web Conference, pp. 30–43. Springer, Berlin (2006)

Rashmi, S., Rangarajan, K.: Rule based visual surveillance system for the retail domain. In: Proceedings of International Conference on Cognition and Recognition, pp. 145–156. Springer, Berlin (2018)

Rusu, R.B., Marton, Z.C., Blodow, N., Dolha, M., Beetz, M.: Towards 3D point cloud based object maps for household environments. Robot. Auton. Syst. **56**(11), 927–941 (2008)

Shen, W., Wang, J., Han, J.: Entity linking with a knowledge base: issues, techniques, and solutions. IEEE Trans. Knowl. Data Eng. **27**(2), 443–460 (2014)

Sirin, E., Hendler, J., Parsia, B.: Semi-automatic composition of web services using semantic descriptions. In: 1st Workshop on Web Services: Modeling, Architecture and Infrastructure, pp. 17–24 (2003)

Sommer, M., Stjepandić, J., Stobrawa, S., von Soden, M.: Automatic generation of digital twin based on scanning and object recognition. Adv. Transdiscip. Eng. **10**, 645–654 (2019)

Staab, S., Studer, R.: Handbook on ontologies. Springer Science & Business Media, Berlin (2010)

Tenorth, M., Beetz, M.: KnowRob—knowledge processing for autonomous personal robots. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4261–4266. IEEE, Piscataway (2009)

Thosar, M., Zug, S., Skaria, A.M., Jain, A.: A review of knowledge bases for service robots in household environments. In: AIC, pp. 98–110 (2018)

Zheng, Y., Yang, S., Cheng, H.: An application framework of digital twin and its case study. J. Ambient. Intell. Humaniz. Comput. **10**(3), 1141–1153 (2019)

A Demand-Response System for Sustainable Manufacturing Using Linked Data and Machine Learning

Hendro Wicaksono, Tina Boroukhian, and Atit Bashyal

Abstract The spread of demand-response (DR) programs in Europe is a slow but steady process to optimize the use of renewable energy in different sectors including manufacturing. A demand-response program promotes changes of electricity consumption patterns at the end consumer side to match the availability of renewable energy sources through price changes or incentives. This research develops a system that aims to engage manufacturing power consumers through price- and incentive-based DR programs. The system works on data from heterogeneous systems at both supply and demand sides, which are linked through a semantic middleware, instead of centralized data integration. An ontology is used as the integration information model of the semantic middleware. This chapter explains the concept of constructing the ontology by utilizing relational database to ontology mapping techniques, reusing existing ontologies such as OpenADR, SSN, SAREF, etc., and applying ontology alignment methods. Machine learning approaches are developed to forecast both the power generated from renewable energy sources and the power demanded by manufacturing consumers based on their processes. The forecasts are the groundworks to calculate the dynamic electricity price introduced for the DR program. This chapter presents different neural network architectures and compares the experiment results. We compare the results of Deep Neural Network (DNN), Long Short-Term Memory Network (LSTM), Convolutional Neural Network (CNN), and Hybrid architectures. This chapter focuses on the initial phase of the research where we focus on the ontology development method and machine learning experiments using power generation datasets.

H. Wicaksono (✉) · T. Boroukhian · A. Bashyal
Jacobs University Bremen, Bremen, Germany
e-mail: h.wicaksono@jacobs-university.de; t.boroukhian@jacobs-university.de; a.bashyal@jacobs-university.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_8

1 Introduction

With the energy transition in Europe toward renewable but irregular resources underway, the need for energy flexibility to balance the weather-dependent energy generation is also increasing. Above all, local energy flexibility, which considers local or own energy generation, plays a decisive role. The Energy Union Framework of the European Union outlines the vision that end users actively participate in the market and benefit from technological progress in the form of cost reductions (Energy Union Package 2015). In order to improve the quality of service, the active participation of end customers is increasingly coming to the foreground. As positive effects, competition will be strengthened, more renewable energy sources will be integrated, and energy networks will be more balanced. This will ensure an efficient operation of the energy systems.

As a prerequisite for this, the majority of energy consumers must be willing to participate actively in demand-response programs (Siano 2014). According to the U.S. Department of Energy (DoE), demand response (DR) refers to Qdr (2006)

> changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.

Demand-response programs can be implemented as price-based or incentive-based programs (Chen and Liu 2017). Price-based programs imply the introduction of dynamic energy tariffs and dynamic demand profiles adapted to changing conditions. Incentive-based programs, on the other hand, require consumers to hand over direct control of their energy systems to third parties for the purpose of modifying the demand profile, if exceptional circumstances in the network should require this.

The modification of consumer demand for energy using methods, such as financial incentives, is known as demand-side management (DSM). DR is a subset of DSM. One objective of DSM is to encourage the consumer to reduce the usage of energy during peak periods or to store energy during off-peak hours using energy storage units. As an application, DSM helps grid operators to provide a balance between renewable generation units such as wind and solar. Due to increasing electricity demand and installation of renewable energy sources, DSM is applied for managing system behavior (Chiu et al. 2012). Therefore, DR systems can be used to attain a balance between the demand and supply in the power grid (Yassine 2016).

The dissemination of demand-response programs in Europe is a slow but steady process. As recently as 2013, European end customers were not very open to DR systems, mainly due to regulatory and legislative barriers. Today, consumers have far more opportunities to participate in demand-response programs in different member states (Bertoldi et al. 2016). Unfortunately, Germany is lagging behind other countries in this area, although demand-response services are being developed in some research. This is due to the doubt of the effectiveness of DR to guarantee energy security and to attract investments (Lehmann et al. 2015; Valdes et al. 2019). Furthermore, critical policy debates on the implementation of DR still exist. The governance needed to relate demand-side response and local energy markets is also still lacking (Kuzemko et al. 2017).

Although some legislative barriers have been dismantled and some aggregators in Europe are installing DR systems in tertiary buildings, the desired effects on the energy grid as a whole have not yet materialized. In Germany, both large industrial consumers and households remain largely unaffected by DR.

Our research project addresses this problem and develops a solution for the participation of industrial end customers in both price- and incentive-based DR programs. The research investigates the influence of such programs on the electrical grid and on the development of energy costs in the manufacturing industry, from different perspectives. This book chapter describes the development of the project's technological solution that allows the implementation of a DR system. The solution uses a semantic middleware that utilizes a linked data approach to connect data sources coming from heterogeneous systems involved at demand and supply sides. This book chapter also explains the development of the forecasting methods to estimate the amount of electricity supply from renewable sources in the network and also the amount of electricity demand from manufacturing companies. The forecasts are used as the foundation to calculate the dynamic electricity tariffs.

Section 2 of this book chapter focuses on the related works in using ontologies or linked data for the integration of data from heterogeneous sources on semantic level and in forecasting methods in energy domain. Section 3 describes the overview of the technical concept that we develop in the project. Then, Sect. 4 shows the method to develop the ontology as the core of the semantic middleware. The methods, the dataset, and some intermediate experiment results of the forecasting are discussed in Sect. 5. Finally, the book chapter ends with conclusions and outlook.

2 Related Works

2.1 Ontologies and Linked Data

Semantic technologies, such as the graph-based data model called Resource Description Framework (RDF) (Lassila et al. 1998) and linked data principles (Bizer et al. 2011) from the Web, facilitate automated information integration. Ontologies provide vocabularies and relation models to make data semantically compatible; thus they can be linked without causing inconsistencies and ambiguities.

Ontology-based information models have been used to facilitate data integration on semantic level, for example to model objects involved in manufacturing and supply chain sustainability (Borsato 2017). In the manufacturing domain, ontologies can model the relationships between products, processes, resources, and environmental factors contributing to energy efficiency (Wicaksono et al. 2014).

Ontologies also facilitate virtual collaborations in assembly automation (Ferrer et al. 2015). The AutomationML ontology is used to enable the communication of heterogeneous tools in engineering environments, such as in manufacturing (Kovalenko and Grangel-Gonzalez 2021).

In the energy management domain, ontologies have been developed to allow a common semantic data model to address interoperability problems among involved building automation systems (Wicaksono et al. 2013, 2010). They are also used to enable the interoperability among Internet-of-Things solutions in general (Alaya et al. 2015; Daniele et al. 2015; Haller et al. 2019) and for energy management applications (Cuenca et al. 2017; Daniele 2021; Kofler et al. 2012). Industry Foundation Classes (IFC) refer to a standard information model allowing exchanges of building data. ifcOWL is the ontology representation of IFC (Pauwels and Terkaj 2016). Researchers also develop ontologies to support the energy transition, for example, to model renewable energy sources (Küçük and Küçük 2018) and smart grid objects (Stap and Daniele 2021).

Open Automated Demand Response (OpenADR) is a standardized information model for sending and receiving DR signals between network operators and consumers (Fernández-Izquierdo et al. 2020). It is implemented as a client–server architecture. Virtual End Nodes (VEN) refer to the clients and Virtual Top Nodes (VTN) are the servers. The DR signals received in the client can trigger preprogrammed actions according to their tasks and thus establish an automated demand response for each resource considered. OpenADR is open source and can therefore be implemented by any user. Since electricity prices generally correlate negatively with electricity availability, consumers can specifically control their consumption away from high price periods and significantly reduce their own energy costs.

In this research, we investigate the scopes of ontologies that are relevant for demand-response systems. We analyze those ontologies based on data requirements at both supply and demand side. Table 1 shows the domain scope of the analyzed ontologies. From the table, it can be seen that no ontology is able to fulfill the data requirements in all domains. Therefore, we develop a linked data approach to interlink those ontologies.

2.2 Machine Learning and Forecasting Methods

The fields of machine learning and artificial intelligence have undergone rapid growth in the last decade, changing the way humans interact with data, especially in the use of data for decision-making. Data-driven decision-making has thus grown in the fields of physical sciences and social sciences, influencing various aspects of daily human life. A huge part of data-driven decision-making entails the need to forecast future behavior of a system or trend of a series based on historical time-series data. Therefore, time-series forecasting has become a crucial method implemented across many domains, where different windows of time series representing past behaviors are utilized to forecast future behaviors.

Table 1 Considered ontologies and their corresponding domain scopes

Ontology	Energy generation	Energy balance	Business data	Environmental factors	Building	Building automation	Industrial process automation	Production planning and control
OpenADR (Fernández-Izquierdo et al. 2020)		✓	✓		✓	✓	✓	
OEMA Ontology Network (Cuenca et al. 2017)	✓		✓	✓	✓			
SAREF4ENER (European Telecommunications Standards Institute 2020)	✓	✓		✓	✓	✓		
ThinkHome (Kofler et al. 2012)	✓				✓	✓		
CIM ontology for Smart Grids (Stap and Daniele 2021)	✓	✓	✓				✓	
OneM2M (Alaya et al. 2015)	✓					✓	✓	
SAREF (Daniele et al. 2015)				✓	✓	✓		
SOSA/SSN (Haller et al. 2019)						✓	✓	
IFC (Pauwels and Terkaj 2016)					✓	✓		
SERENE (Wicaksono 2016)					✓	✓	✓	✓
AutomationML (Kovalenko et al. 2018)						✓		✓
PPR model (Ferrer et al. 2015)								✓
MASON (Lemaignan et al. 2006)								✓
SERUM/ KnoHolEM ontology (Wicaksono et al. 2013)				✓	✓	✓		
OntoWind (Küçük and Küçük 2018)	✓	✓						

Existing methods for time-series forecasting can roughly be mapped into two categories, i.e., classical methods that focus on learning linear relationships and machine learning/deep learning methods that generalize complex non-linear relationships. Classical time-series models such as Exponential Smoothing, Autoregression, and Moving Average have easy interpretability and strong theoretical assurance. Simple variations of classical models, such as Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA), have also been established as reliable methods for time-series forecasting. More complex extensions of these methods, which include variations with integration of seasonality: Seasonal Autoregressive Integrated Moving Average (SARIMA), and integration of vectorized multiple inputs: Vector Autoregression (VAR) and Vector Autoregression Moving Average (VARMA), further provide flexibility in adding complex relations in the forecasting and models. Novel extensions of these methods also support handling missing data and multiple data types (Seeger et al. 2016).

Machine learning/deep learning methods aim to learn a non-linear function mapping of stochastic historic input data to a predicted/forecasted output value. Methods such as Support Vector Regression (SVR) (Zafirakis et al. 2019; Lahouar and Slama 2017), Random Forest (RF) (Li and Shi 2010; Sun et al. 2018), and Artificial Neural Network (ANN) (Jiao et al. 2018; Olaofe and Folly 2013) are the simplest existing techniques that have shown abilities in learning these non-linear functions. Extending the ANN models to have recurrent hidden layers (RNN) that capture and preserve the relationship of a new input with previous inputs achieved ground-breaking performance in sequence learning. Improved variants of RNN, i.e., Long Short-Term Memory (LSTM) and Grated Recurrent Unit (GRU), (Liu et al. 2019; Niu et al. 2020) have also shown great potential in learning complex relationships present in stochastic historic input data for time-series forecasting. As one of the most popular deep learning methods, Convolutional Neuron Network (CNN) has not only been utilized to extract visual features in problems of computer vision and natural language processing (NLP), but is also used in feature selectors and predictors for wind speed/power forecasting. The work of Liu et al. (2019) introduced a 1D CNN-based forecasting model for direct multi-steps ahead prediction for short-term wind speed data, which outperformed other classical and baseline ANN methods. Wang et al. (2019) used 2D CNN to efficiently extract the non-linear features of the raw wind power data, which shows better accuracy in probabilistic forecasting.

Alongside the development of forecasting methods, research works have increased in forecasting both electricity generation (Wang et al. 2019) and consumption for industrial and household consumers (Hong et al. 2020). Most of the recent research works on electricity generation focus on production of green energy from wind and solar plants (Liu et al. 2019; Wang et al. 2019). On the consumption side, research is driven by optimizing energy efficiency, which also requires constant monitoring of various consumption indicators and identifying factors that affect them in real time (Hong et al. 2020). Weather conditions are identified as the main factor determining the production of green energy and the demand for electricity. Additional independent factors, such as holidays, operational

Table 2 Overview of deep learning forecast methods from related works

Algorithm name/source	Multivariate input	Input series frequency	Forecast horizon
SAE-BP (Huang et al. 2017)	×	15 min	2 h
EHS-SVR (Zafirakis et al. 2019)	✓	15 min	3 h
RF-AR, SVR-AR (Lahouar and Slama 2017)	✓	1 h	1–2 h
RF (Sun et al. 2018)	✓	15 min	15 min–24 h
BPNN (Li and Shi 2010)	✓	1 h	24 h
BPNN (Jiao et al. 2018)	×	1 h	1 h
LSTM (Olaofe and Folly 2013)	✓	15 min	2 days
Attention-GRU (Niu et al. 2020)	✓	1 h	3 h
CNN-LSTM (Liu et al. 2019)	×	10 min	30 min
Deep CNN (Wang et al. 2019)	×	5 min	8 h

characteristics of buildings, and indicators of living standards, are also identified to influence the demand (Hong et al. 2020; Kalimoldayev et al. 2020). These advancements in forecast form the basis to drive research work on dynamic pricing of electricity and implementing DR systems (Kalimoldayev et al. 2020). The influence of DR systems has also been evaluated with the use of deep learning methods. Micro-grid demand-response systems have shown high potential in not only helping consumers reduce their energy bills, but also help in improving grid stability and reliability (Shojaeighadikolaei et al. 2021).

Therefore, forecasting is seen as a fundamental aspect of planning and management, and the design of a dynamic price system requires forecasting of future supply and demand (Dutta and Mitra 2017). Further, scheduling of actual consumption would also require forecasts of future prices. Therefore, it is identifiable that an initial step toward building a price system is to have a strong forecasting method. And the potential that deep learning shows in mapping relations in data for forecasting makes exploring the deep learning method an important initial step. Thus, experimental works are done in our research to create baseline forecasting methods using deep learning. Details of the experimental works are discussed in Sect. 5. Table 2 shows different existing research works related to forecasting methods, especially in terms of renewable energy. We reviewed the works by analyzing the algorithms used or developed, whether the input data is multivariate or not, the frequency of the input data points, and the forecast horizon.

3 Overview of the Concept

The architecture of the solution is based on the reference architecture for the smart grid developed by the EU Smart Grid Coordination Group/Reference Architecture Working Group (SG-CG/RA) (Gottschalk et al. 2017) (see Fig. 1). The core of

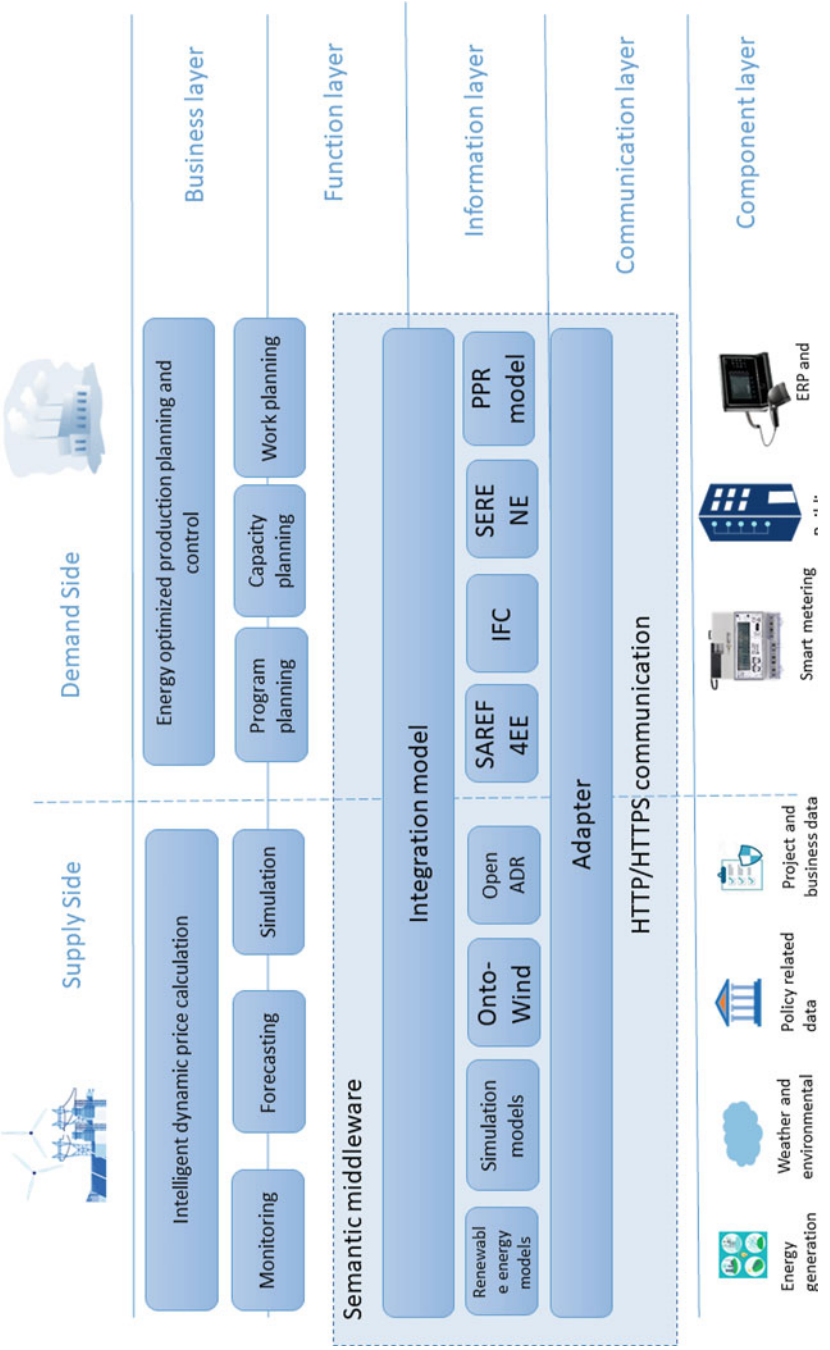


Fig. 1 Overview of the solution

the architecture is the semantic middleware. This is a software module with an ontology for the integration of heterogeneous information models, which are defined by standards from different subject areas (e.g., OpenADR, IFC, SAREF, PPR model). A method for mapping between the DR communication standards and the ontologies is developed for the semantic middleware to enable automatic data routing. Using the linked data approach, the heterogeneous information models shown in Table 1 are semantically linked. The adapter serves as a communication interface between various systems supplying the data. The adapter contains an HTTP/HTTPS interface and a software module for data format conversion. The applications of the energy supply and demand sides use the semantic middleware to automatically collect the data required for their functionalities. An intelligent mapping and routing algorithm is developed to capture the appropriate data from the right data sources.

In our research, data are collected from heterogeneous sources. Those data include power generation, power consumption, weather, customer, balancing group, industrial automation, and manufacturing process data. We develop a continuous intelligent analysis based on this data, i.e., descriptive data analysis for monitoring, predictive analysis to generate the forecast model of electricity supply and demand, prescriptive analysis through simulation and optimization to determine the dynamic electricity prices, as well as the visualization of the data to support the analysis. This continuous data analysis makes the determination of the dynamic electricity prices more precise and transparent. This will happen on the utility side. The end-to-end intelligent data analyses are also used on the consumption side. Load profiles of the production processes are recorded based on electricity demand measurements at the systems involved in the production process, and connections between production orders and the associated electricity demand are created using machine learning methods. These findings are then used to generate electricity demand forecasts depending on the production plan. Heuristics are developed to enable electricity cost-optimized production planning and control. DR requirements are also explicitly considered.

The concepts will be examined in the context of two use cases, i.e., at a rubber component manufacturer and at a manufacturing company having a high degree of automation. Two different perspectives are considered in each use case: the technical and the economic feasibility. The use case at the rubber component manufacturer focuses on the development and integration of the hardware required to implement demand-response requirements. The use case of the second company builds on existing sensors and develops simulation models for the optimization of production by exchanging energy data with the market. Those companies have different requirements for market mechanisms and negotiation strategies, which are analyzed within the scope of our research.

The following sections focus on the development of the ontology for the semantic middleware and the methods to forecast power generation and consumption as the initial phase of end-to-end intelligent data analysis. Since the project is still in the early phase, in this book chapter, we only concentrate on the development and experiments of forecasting for wind and solar power generation. However, the developed methods can be applied to the forecast of power generation from other types of renewable sources and also to the forecast of power consumption.

4 Ontology Development

4.1 Ontology Construction Methodology

In this research, to represent a sharable and reusable knowledge as a set of concepts within the domain and to describe their relationships, ontologies as a part of the W3C standards stack for the Semantic Web are applied. They provide a link between different pieces of information from different domains. Ontologies are also used to enhance knowledge and data exchange in these domains.

To achieve this goal, we adapt an ontology construction method toward data integration through linked data developed for building energy management domain (McGlinn et al. 2016). The adaptation of the method is shown in Fig. 2. First, we define data requirements and the mapping of those data to appropriate ontologies. Then, the method of systematic review is applied with the purpose of identifying and extracting ontologies and analyzing the potential links between them in the following specific fields: demand response, sensor networks, renewable energy, and manufacturing. The following sections explain the steps carried out in the method in detail.

Step 1: Data Requirements

Step 1 analyzes the specific data requirements for each process in the system in more detail. The purpose of this step is to understand the exact structure of the data required to meet the use cases or functionalities of the system. All data values that are required must be captured and described. This involves structuring the data as concepts and properties.

Step 2: Ontology Domains Finding and Extracting

The purpose of this step is to provide a quick reference to ontologies related to demand-response systems and select the best of them for a particular domain (step 2a). These ontology domains are extracted and presented here. The next step is to explore and find potential links between ontologies in the abovementioned categories (step 2b and 2c).

Step 3: Develop Ontology

Step 3 is concerned with the development of models for meeting the data requirements, which are not currently supported by any existing ontologies or standards.

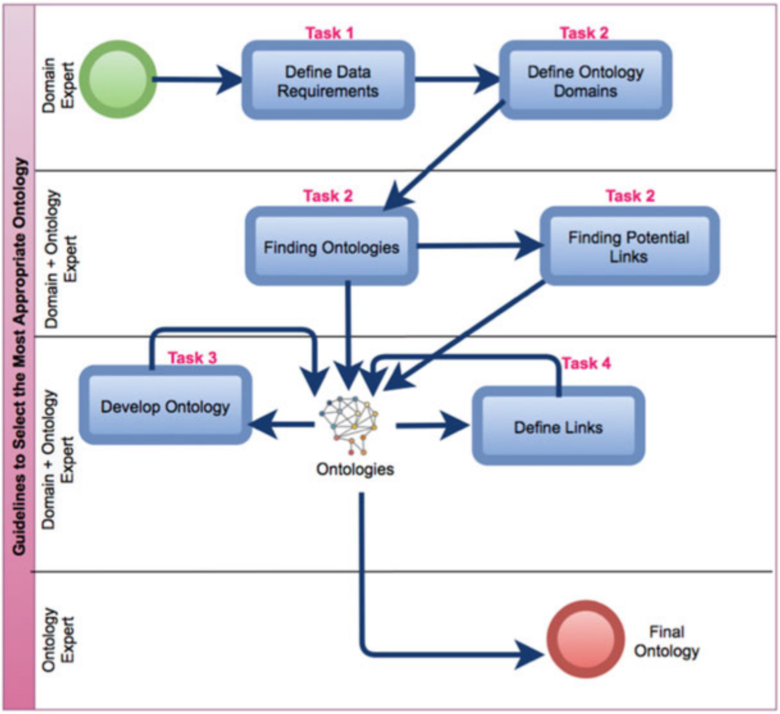


Fig. 2 Ontology development

Step 4: Define Links

Step 4 is concerned with the definition of links between ontologies, since multiple ontologies are required to meet the data requirements. At this stage, the mappings and alignments must be identified and formalized.

4.2 Database to Ontology Mapping

In our project, we collect data from different sources. Those data can be in relational databases, CSV files, or XML formats. A relational database is commonly used in modern applications for storing and querying data and enables the organization of data points and identifies their relationships. Since an ontology provides formal semantics for the data explicitly, to describe the semantics of data stored in a database, the database schemas should be converted into semantically equivalent ontologies. Mapping rules are applied for converting relational databases into ontology. The database to ontology mapping is located in the adapter (see Fig. 1). We develop the database to ontology mapping by considering the works described in Mahria et al. (2021), Hazber et al. (2015).

4.3 Ontology Alignment

Since we consider multiple ontologies to fulfill the data requirements, we develop ontology alignment methods to interlink the ontologies. An ontology alignment is an ontology matching process to provide a set of correspondences between semantically related ontology concepts for two input ontologies (one source ontology and one target ontology). These semantic relationships are called mappings. To determine the mapping, we find similarities between the entities of semantically related ontologies. As an important task, ontology alignment allows the joint consideration of resources described by different ontologies. Ontology alignment is used to solve different problems of semantic heterogeneity in the integration and sharing of information. We apply multiple ontology alignment approaches that are based on lexical, structural, extensional, and semantic techniques (Ouali et al. 2019; Xie et al. 2016).

A lexical method is based on the comparison of terms, strings, or texts. But structural methods calculate the similarity between two entities by exploiting structural information within semantic or syntactic links, which form a hierarchy. Extensional methods infer the similarity between two entities by analyzing their extensions (i.e., their instances). Finally, semantic methods are based on the external ontologies. In Nguyen and Conrad (2015) and Essayeh and Abed (2015), the authors develop hybrid methods (combining structural and semantic).

Because of the extensive applications of ontology alignment, it has been widely studied in many research works (Mohammadi 2019; Mohammadi et al. 2018a,b; Mohammadi and Rezaei 2020; Zhou et al. 2018). In these works, the authors also consider a considerable number of alignment systems. In our research, the ontology alignment generates the integration model shown in Fig. 1.

4.4 Ontology Candidates and Potential Mapping

In this section, we provide a quick reference to ontologies related to demand response and select the best of them in the following domains: demand response, renewable energy, sensor, and manufacturing. The next step is to explore the potential links between ontologies. We interlink those ontologies to prevent information duplication, to establish a common vocabulary, and to make less development effort. We consider methods `SubClassOf` and `EquivalentClass` to link a concept in the source ontology and a concept in the target ontology. In the `SubClassOf` approach, concepts and properties of source ontology are a part of the target ontology. In `EquivalentClass`, the linked concepts refer to the same meaning. Sometimes, to be more flexible, we may define "new concepts" or "new properties" for the source and target ontologies (Haase and Motik 2005).

The lists of ontologies that we considered in this context are OpenADR, OEMA, SAREF, OntoWind, SSN/SOSA, and Mason. We elaborate the potential links

Table 3 Potential links between different ontologies: OpenADR

Source ontology	Target ontology	Link type	Description
`oadr:Resource`	`mason:Resource`	`EquivalentClass`	Same concepts
`oadr:Item`	`saref:Property`	SubClassOf	This link is specified in Fernández-Izquierdo et al. (2020)
`oema(PAO):Organisation`	`openadr:Resource`	`SubClassOf`	Resource is the entity in the DR programs, and organization can join to the DR programs

between the abovementioned ontologies and show the linking method in Tables 3, 4, 5, 6, 7 and 8.

5 Forecast Methods

As the literature discussed in the previous section shows, artificial neural networks and deep learning have recently been popular in mapping stochastic historic input data to a forecast output value. The initial experiments done in our research are to develop simple deep learning networks and evaluate their prediction capabilities. In order to do so, four neural network architectures, Deep Neural Network (DNN), Long Short-Term Memory Network (LSTM), Convolutional Neural Network (CNN), and a Hybrid architecture combining the three techniques, are developed. We conduct experiments to forecast both solar and wind energy generation for both short-term (1 day) and long-term (monthly) horizons. The details of the experiments and their results are discussed in the following order. The datasets and data preprocessing techniques are discussed first, along with exploratory data analysis. This is followed by discussions on the neural network architectures. Finally, the performance of the different architectures is discussed.

5.1 Dataset Description, Data Preprocessing, and Exploratory Data Analysis

Data used for experimental purposes are historical wind and solar power generation data, from February 2019 until December 2020, collected by a local utility company. The wind power generation data are from three different wind turbines and include variables such as the wind speed, wind direction, and temperature around the turbines. Similarly, the solar power generation data are also from three different solar plants and include other variables such as global radiation and temperature around the solar plants. The dataset preprocessing is done with simple techniques to identify abnormal/anomaly values, drop them, and finally impute missing values using linear interpolation.

Table 4 Potential links between different ontologies: OEMA

Source ontology	Target ontology	Link type	Description
`oema:Sensor (in Energy and Equipment)`	`ssn/sosa:Sensor`	`EquivalentClass`	Same concepts
`oema:WindSpeed(in Infrastructur)`	`ontowind:WindSpeed`	`EquivalentClass`	Same concepts
`mason: HumanResource`	`oema: OrganisationMember (in Person and Organisation)`	`SubClassOf`	HumanResource as a manufacturing concept can be a member of an Organization

Table 5 Potential links between different ontologies: SAREF

Source ontology	Target ontology	Link type	Description
`Saref:Function`	`ssn/sosa:Procedure`	`EquivalentClass`	In SSN, Procedure class has properties ssn:hasInput and ssn:hasOutput that implement function logic.
`Saref:Temperature`	`ontowind: Temperature`	`EquivalentClass`	Same concepts

Table 6 Potential links between different ontologies: Mason

Source ontology	Target ontology	Link type	Description
`mason: Machine-tool`	`saref: Device`	`SubClassOf`	Machine-tool with concept of manufacturing can be a Device that is a tangible object designed to accomplish a particular task.
`mason: Scheduling`	`openadr:Schedule`	`EquivalentClass`	Same concepts

Table 7 Potential links between different ontologies: OntoWind

Source ontology	Target ontology	Link type	Description
`ontowind:Sensor`	`ssn/sosa:Sensor`	`EquivalentClass`	Same concepts
`ontowind:Pressure`	`saref:Pressure`	`EquivalentClass`	Same concepts
`ontowind:Generator`	`saref:Generator`	`EquivalentClass`	Same concepts

Table 8 Potential links between different ontologies: SSN/SOSA

Source ontology	Target ontology	Link type	Description
`ssn/sosa: FeatureOfInterest`	`saref: FeatureOfInterest`	`EquivalentClass`	Same concepts
`ssn/sosa:Procedure`	`saref:Function`	`EquivalentClass`	In SSN, Procedure class has properties ssn:hasInput and ssn:hasOutput that implement function logic

As a next step to understand the datasets, calculation of correlations between all variables is necessary. Initially, no correlation is seen between the generated power and the wind direction. This problem generally appears because the coordinate in which wind direction is measured is circular, opposed to the linear coordinate used to measure wind speed and power. In order to extract a new feature that encompasses the effect of wind direction and wind speed, the following trigonometric transformation is applied to the initial features:

$$\text{Output}_d = \text{wind\_speed} \times \cos^2(d - \text{wind\_direction}). \tag{1}$$

In the equation, d is an offset direction such that: wind direction values closer to the offset direction result in higher production. Two offset directions are chosen based on the apparent relation seen between the direction feature and the generated power.

The histogram shown in Fig. 3 has power generated on the y-axis plotted with the wind direction (0° as North) for each month of the year. In each plot, we can

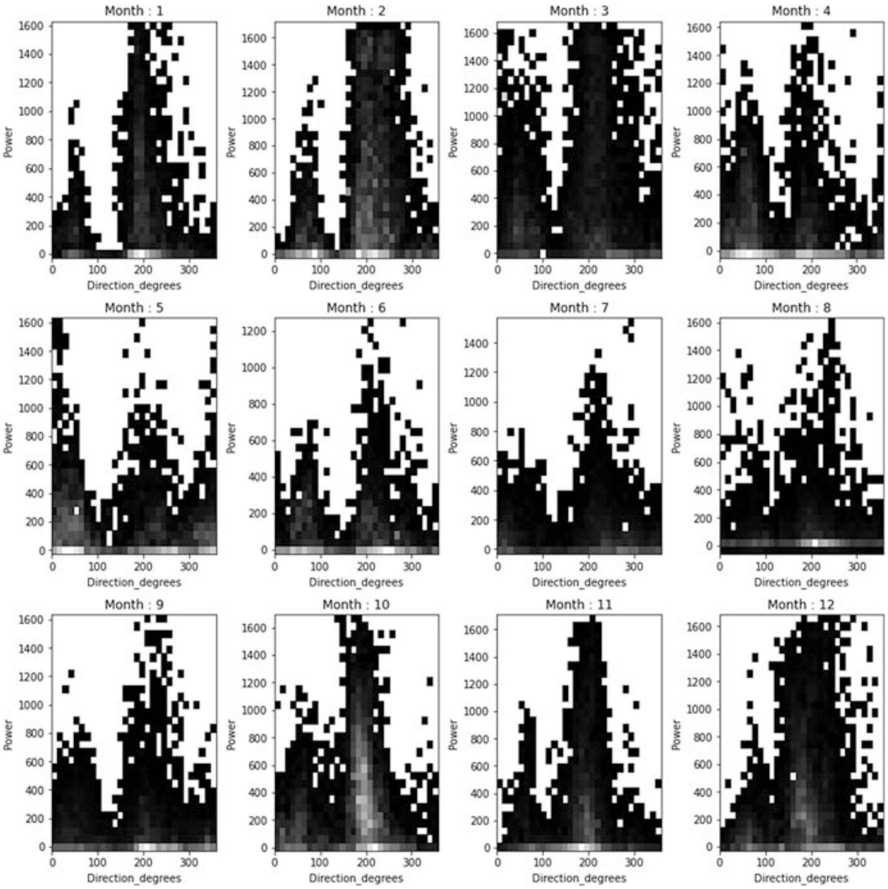


Fig. 3 Histogram of power against wind direction

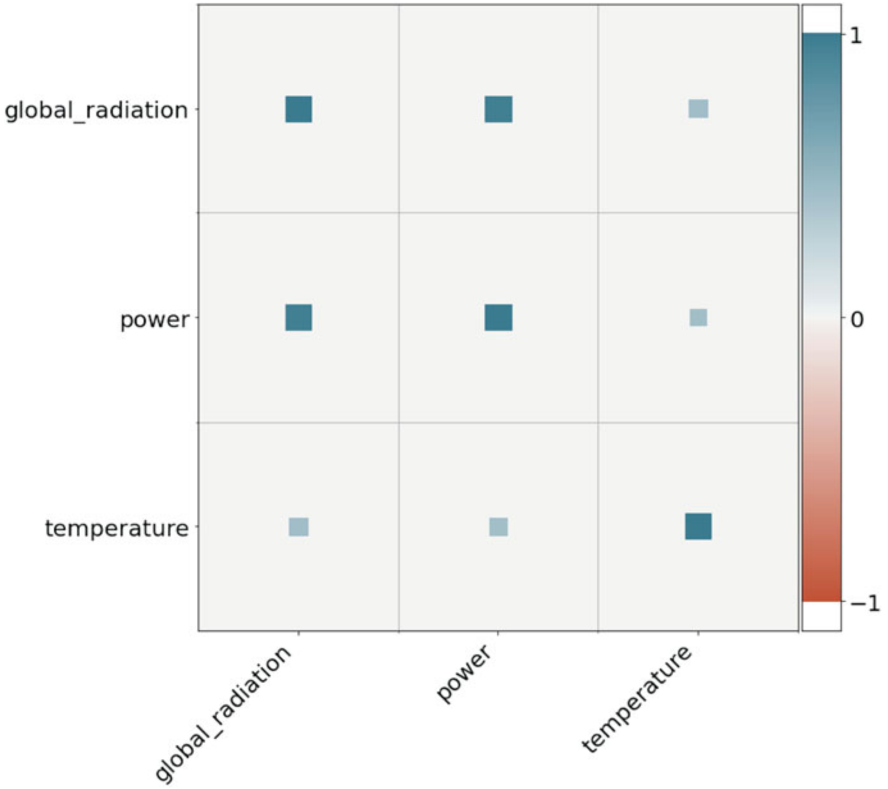


Fig. 4 Correlation plot: features used for solar power prediction

see two distinct clusters form. The first cluster is centered around 45–50° North, and the second at 200–225° North. It is also slightly noticeable that the closer the wind direction to these axes, the higher wind energy production is seen (bright spots in the histogram). We therefore estimate that the ideal directions for the wind turbine output are around 50 and 200° North, and these are the ideal offset angles to choose. The solar dataset shows good correlations of the initial features with power generated. Figures 4 and 5 show correlation plots of the variables used to train neural network models for wind and solar power forecast.

As mentioned earlier, the initial experiments are conducted with Deep Neural Network (DNN), Long Short-Term Memory Network (LSTM), Convolutional Neural Network (CNN), and Hybrid architecture. Each training step allows the network to learn from historical values with a window of 18 time steps in order to predict power generation for the next single time step. One year's half power generation data are used for training and the other half for evaluation of the models. Table 9 below summarizes the network architecture for each of the methods.

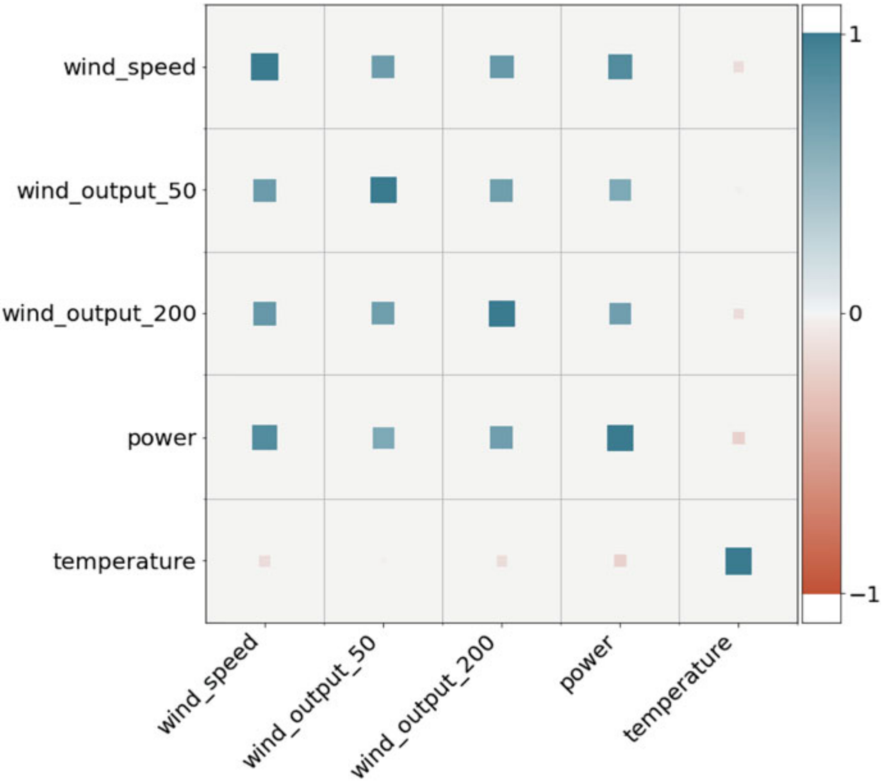


Fig. 5 Correlation plot: features used for wind power prediction

5.2 *Experiment Results*

Each model is trained to minimize the mean-squared error between the predicted value and the actual value. Therefore, the mean-squared error (MSE) is used to evaluate the different neural network architectures. Along with the mean-squared error, the mean absolute percentage error (MAPE) of the different models is compared. Tables 10 and 11 summarize the metrics for short-term and long-term forecasting of wind and solar power generation.

The results show that the Hybrid model outperforms all other models and significantly improves prediction in terms of MSE for daily and monthly predictions. Theoretically, the Hybrid model benefits from the integration of the LSTM and CNN layers. The LSTM layer can learn the temporal differences in the data (dependence on previous values), and the CNN layer is able to learn the spatial relations (correlations between input features). Both the short-term predictions and long-term predictions of the Hybrid model show high flexibility in terms of predicting but still tend to overestimate the forecasted values resulting in higher MAPE values.

Table 9 Overview of deep learning architectures

Model	Parameters	Comments
DNN	(Input)64|32,16(hidden)|1(output)	Input does not include historical values of power. Multivariate model including weather condition and temperature as input features
LSTM	(Input)64|100(encoder)|100(decoder)|1(output)	Input includes historical values of power along with weather condition and temperature as input features
CNN	(Input)64|(2-D conv)|1(output)	Time distributed 2-D Convolutional Network, with 2 convolutional layers. The layers have 8 and 4 filter channels, respectively. Each filter channel has kernels of sizes 4 and 2, respectively. Input includes historical power value, weather condition, and temperature as input features
DNN-LSTM-CNN	(Input)64|CNN-LSTM-DNN|1(output)	A hybrid model that combines architectures mentioned above. Network Architecture is set up in the following order, input layer-CNN layer-LSTM layer-DNN layer. Input includes historical power value, weather condition, and temperature as input features

Table 10 Model evaluations for short-term forecasting. The best results according to the evaluation metrics are highlighted in boldface

	MSE		MAPE	
Model	Wind	Solar	Wind	Solar
DNN	24e3	58e3	141.2%	71.2%
LSTM	17e4	46e6	**73.4%**	**48.6%**
CNN	18e4	13e3	103.8%	51.1%
Hybrid	**10e3**	**12e3**	90.1%	59.8%

Table 11 Model evaluations for long-term forecasting. The best results according to the evaluation metrics are highlighted in boldface

	MSE		MAPE	
Model	Wind	Solar	Wind	Solar
DNN	50e3	78e3	10.6%	7.0%
LSTM	71e4	48e6	**23.0%**	**6.8%**
CNN	73e4	83e3	50.8%	16.7%
Hybrid	**46e3**	**15e3**	72.8%	11.4%

Figures 6, 7, 8, and 9 show the plots of short-term and long-term forecasting of wind and solar power generation using the Hybrid model.

The LSTM model stands out in terms of MAPE as it tends to under-forecast predicted values of both wind and solar predictions. The calculation of MAPE favors models that under-forecast and penalize over-forecast (where predicted values are

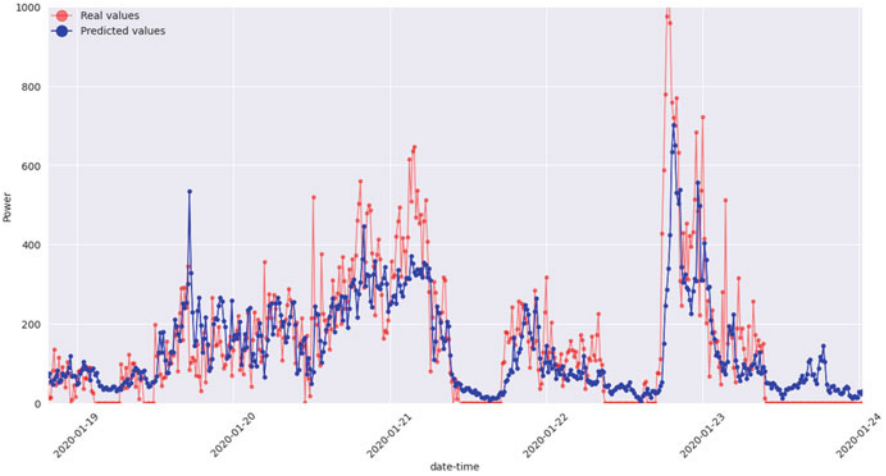


Fig. 6 Short-term forecasting of wind power generation using Hybrid model

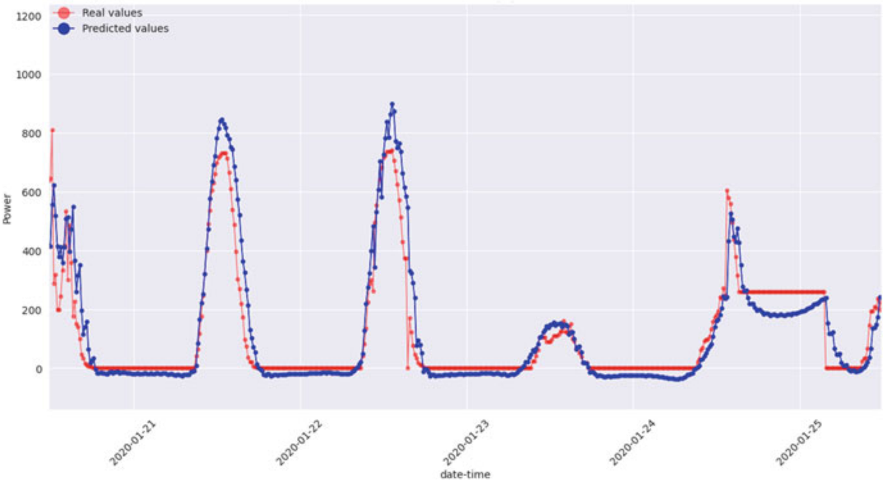


Fig. 7 Short-term forecasting of solar power generation using Hybrid model

higher than actual values). This can be clearly seen in the evaluation plots of the LSTM model in Figs. 10, 11, 12, and 13.

6 Conclusions and Outlook

This book chapter focuses on the development of an IT system that enables the implementation of a demand-response system for manufacturing power consumers. The system architecture follows the functionalities of the reference architecture

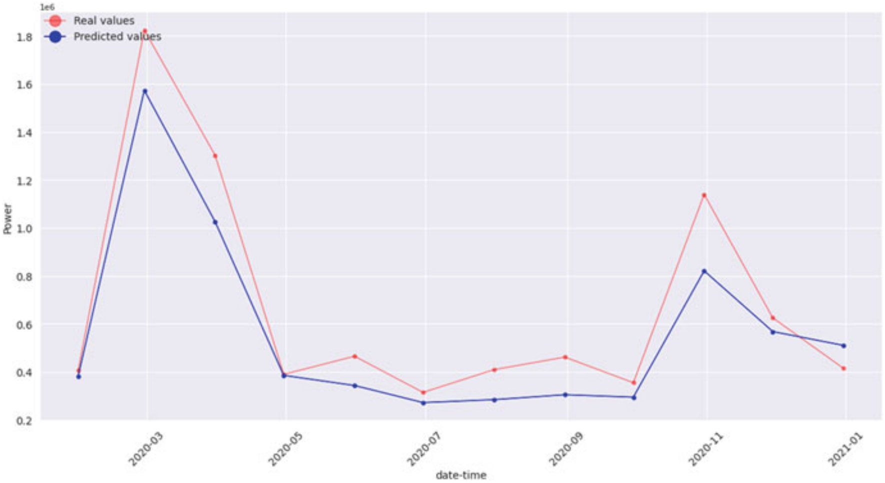


Fig. 8 Long-term forecasting of wind power generation using Hybrid model

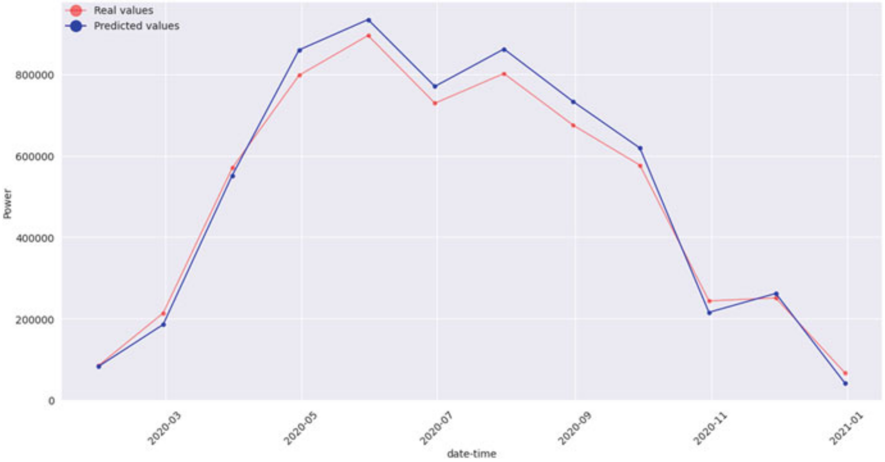


Fig. 9 Long-term forecasting of solar power generation using Hybrid model

developed by EU Smart Grid Coordination Group/Reference Architecture Working Group (SG-CG/RA). The system employs ontologies as the information model to enable the interoperability of heterogeneous systems involved in demand-response programs at both supply and demand sides. We develop an ontology construction method that allows reusing and interlinking of existing ontologies and information model standards, such as SAREF, OpenADR, IFC, Mason, etc. By doing this, many systems that use existing information model standards will be compatible and able to communicate with our system.

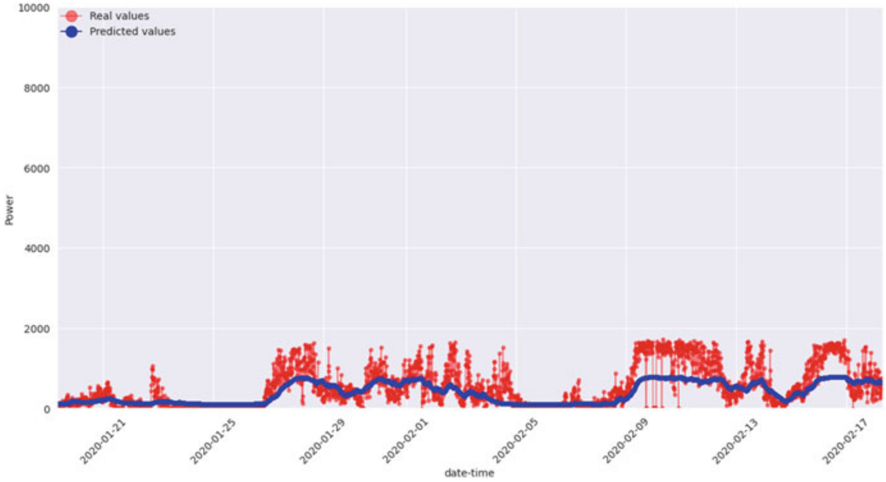


Fig. 10 Short-term forecasting of wind power using LSTM

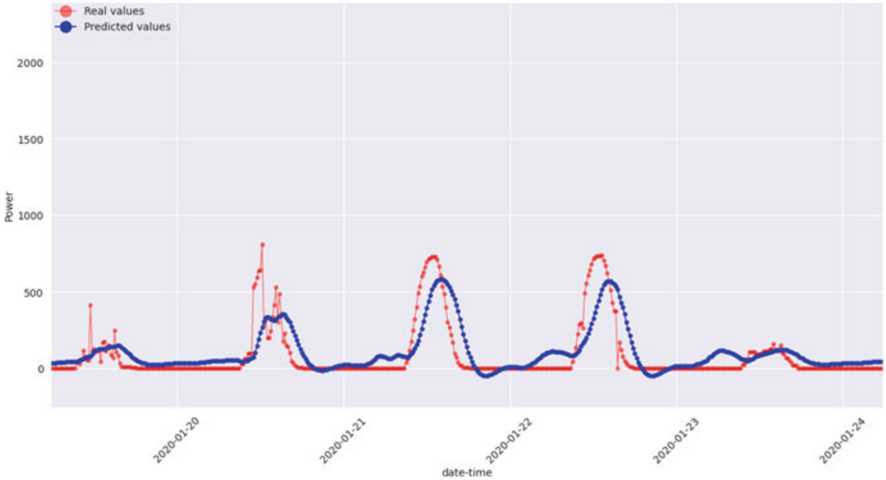


Fig. 11 Short-term forecasting of solar power using LSTM

This book chapter also discusses the development of forecast models as a component of the demand-response system. The aim of the forecast models is to allow for forecasting of the amount of power generated by renewable energy sources, such as wind and solar, and the amount of power consumed in the manufacturing companies. Therefore, we are able to calculate the dynamic electricity tariffs. We develop four different neural network architectures and conduct experiments using solar and wind power generation datasets for short-term and long-term forecasting. The experimental results show that the hybrid and LSTM architectures perform best.

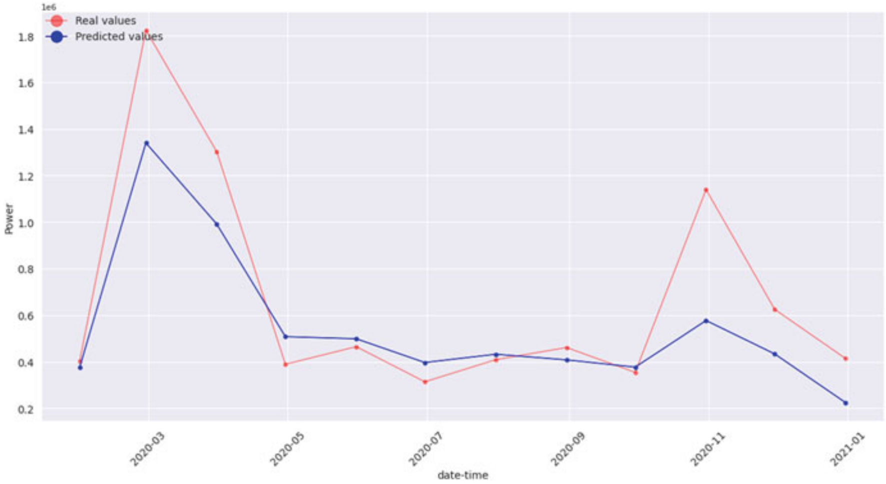


Fig. 12 Long-term forecasting of wind power using LSTM

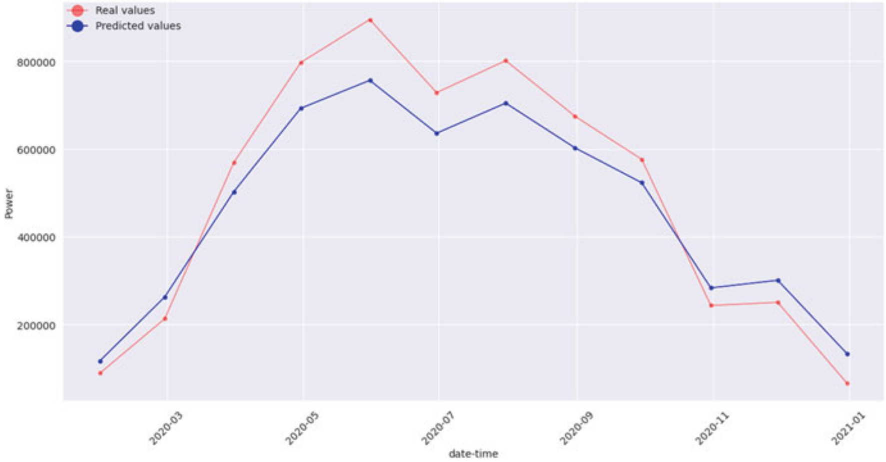


Fig. 13 Long-term forecasting of solar power using LSTM

The research described in this book chapter is still in the early phase. The next steps are to apply the neural network models to forecast the power consumption in manufacturing companies. We will collect the power consumption and manufacturing process data from two manufacturing companies, i.e., a rubber component manufacturer and a manufacturing subcontractor having high-degree automation. Then, we will validate the models using those data. Finally, we will develop a method to calculate dynamic electricity tariffs based on the forecast models.

Acknowledgments This research is funded by German Federal Ministry of Economics and Technology (BMWi). We thank all project members who have contributed in preliminary works and discussions for this research.

References

European Telecommunications Standards Institute (ETSI). https://saref.etsi.org/saref4ener/v1.1.2/. Accessed 06 April 2020

Energy Union Package: A framework strategy for a resilient energy union with a forward-looking climate change policy. Technical report. European Commission, Brussels (2015)

Zafirakis, D., Tzanes, G., Kaldellis, J.K.: Forecasting of Wind Power Generation with the Use of Artificial Neural Networks and Support Vector Regression Models. Energy Procedia **159**, 509–514 (2019). https://doi.org/10.1016/j.egypro.2018.12.007. https://www.sciencedirect.com/science/article/pii/S187661021831302X

Alaya, M.B., Medjiah, S., Monteil, T., Drira, K.: Toward semantic interoperability in oneM2M architecture. IEEE Commun. Mag. **53**(12), 35–41 (2015)

Bertoldi, P., Zancanella, P., Boza-Kiss, B., et al.: Demand response status in EU member states. In: EUR 27998 EN, pp. 1–140 (2016)

Bizer, C., Heath, T., Berners-Lee, T.: Linked data: the story so far. In: Semantic Services, Interoperability and Web Applications: Emerging Concepts, pp. 205–227. IGI Global, Pennsylvania (2011)

Borsato, M.: An energy efficiency focused semantic information model for manufactured assemblies. J. Clean. Prod. **140**, 1626–1643 (2017). https://doi.org/10.1016/j.jclepro.2016.09.185. https://www.sciencedirect.com/science/article/pii/S0959652616315268

Chen, S., Liu, C.C.: From demand response to transactive energy: state of the art. J. Mod. Power Syst. Clean Energy **5**(1), 10–19 (2017)

Chiu, W.Y., Sun, H., Poor, H.V.: Demand-side energy storage system management in smart grid. In: 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm), pp. 73–78. IEEE, Piscataway (2012)

Cuenca, J., Larrinaga, F., Curry, E.: A unified semantic ontology for energy management applications. In: WSP/WOMoCoE@ ISWC, pp. 86–97 (2017)

Daniele, L.: SAREF4EE: the extension of SAREF for EEbus and Energy@Home. https://ontology.tno.nl/saref4ee/. Accessed 07 Feb 2021

Daniele, L., den Hartog, F., Roes, J.: Created in close interaction with the industry: the smart appliances reference (SAREF) ontology. In: International Workshop Formal Ontologies Meet Industries, pp. 100–112. Springer, Berlin (2015)

Dutta, G., Mitra, K.: A literature review on dynamic pricing of electricity. J. Oper. Res. Soc. **68**(10), 1131–1145 (2017)

Essayeh, A., Abed, M.: Towards ontology matching based system through terminological, structural and semantic level. Proc. Comput. Sci. **60**, 403–412 (2015)

Fernández-Izquierdo, A., Cimmino, A., Patsonakis, C., Tsolakis, A.C., García-Castro, R., Ioannidis, D., Tzovaras, D.: OpenADR ontology: semantic enrichment of demand response strategies in smart grids. In: 2020 International Conference on Smart Energy Systems and Technologies (SEST), pp. 1–6. IEEE, Piscataway (2020)

Ferrer, B.R., Ahmad, B., Lobov, A., Vera, D.A., Lastra, J.L.M., Harrison, R.: An approach for knowledge-driven product, process and resource mappings for assembly automation. In: 2015 IEEE International Conference on Automation Science and Engineering (CASE), pp. 1104–1109. IEEE, Piscataway (2015)

Gottschalk, M., Uslar, M., Delfs, C.: The Use Case and Smart Grid Architecture Model Approach: The IEC 62559-2 Use Case Template and the SGAM Applied in Various Domains. Springer, Berlin (2017)

Haase, P., Motik, B.: A mapping system for the integration of OWL-DL ontologies. In: Proceedings of the First International Workshop on Interoperability of Heterogeneous Information Systems, pp. 9–16 (2005)

Haller, A., Janowicz, K., Cox, S.J., Lefrançois, M., Taylor, K., Le Phuoc, D., Lieberman, J., García-Castro, R., Atkinson, R., Stadler, C.: The modular SSN ontology: a joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation. Semantic Web **10**(1), 9–32 (2019)

Hazber, M.A., Li, R., Zhang, Y., Xu, G.: An approach for mapping relational database into ontology. In: 2015 12th Web Information System and Application Conference (WISA), pp. 120–125. IEEE, Piscataway (2015)

Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D., Zareipour, H.: Energy forecasting: a review and outlook. IEEE Open Access J Power Energy **7**, 376–388 (2020). http://doi:10.1109/OAJPE.2020.3029979

Huang, C.M., Kuo, C.J., Huang, Y.C.: Short-term wind power forecasting and uncertainty analysis using a hybrid intelligent method. IET Renew. Power Gener. **11**(5), 678–687 (2017)

Jiao, R., Huang, X., Ma, X., Han, L., Tian, W.: A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting. IEEE Access **6**, 17851–17858 (2018)

Kalimoldayev, M., Drozdenko, A., Koplyk, I., Marinich, T., Abdildayeva, A., Zhukabayeva, T.: Analysis of modern approaches for the prediction of electric energy consumption. Open Eng. **10**(1), 350–361 (2020)

Kofler, M.J., Reinisch, C., Kastner, W.: A semantic representation of energy-related information in future smart homes. Energy Build. **47**, 169–179 (2012)

Kovalenko, O., Grangel-Gonzalez, I.: AutomationML ontology. http://i40.semantic-interoperability.org/automationml/Documentation/index.html. Accessed 07 Mar 2021

Kovalenko, O., Grangel-González, I., Sabou, M., Lüder, A., Biffl, S., Auer, S., Vidal, M.E.: AutomationML ontology: Modeling Cyber-Physical Systems for Industry 4.0. IOS Press Journal, Amsterdam (2018)

Küçük, D., Küçük, D.: Ontowind: an improved and extended wind energy ontology. arXiv preprint arXiv:1803.02808 (2018)

Kuzemko, C., Mitchell, C., Lockwood, M., Hoggett, R.: Policies, politics and demand side innovations: the untold story of Germany's energy transition. Energy Res. Soc. Sci. **28**, 58–67 (2017). https://doi.org/10.1016/j.erss.2017.03.013. https://www.sciencedirect.com/science/article/pii/S221462961730083X

Lahouar, A., Slama, J.B.H.: Hour-ahead wind power forecast based on random forests. Renew. Energy **109**, 529–541 (2017)

Lassila, O., Swick, R.R., et al.: Resource Description Framework (RDF) model and syntax specification (1998). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.6030

Lehmann, P., Brandt, R., Gawel, E., Heim, S., Korte, K., Löschel, A., Massier, P., Reeg, M., Schober, D., Wassermann, S.: Capacity payments to secure electricity supply? On the future of Germany's power market design. Energy Sustain. Soc. **5**(1), 1–7 (2015)

Lemaignan, S., Siadat, A., Dantan, J.Y., Semenenko, A.: MASON: a proposal for an ontology of manufacturing domain. In: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06), pp. 195–200. IEEE, Piscataway (2006)

Li, G., Shi, J.: On comparing three artificial neural networks for wind speed forecasting. Appl. Energy **87**(7), 2313–2320 (2010)

Liu, H., Mi, X., Li, Y., Duan, Z., Xu, Y.: Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional gated recurrent unit network and support vector regression. Renew. Energy **143**, 842–854 (2019)

Mahria, B.B., Chaker, I., Zahi, A.: A novel approach for learning ontology from relational database: from the construction to the evaluation. J. Big Data **8**(1), 1–22 (2021)

McGlinn, K., Weise, M., Wicaksono, H.: Towards a shared use case repository—the SWIMing initiative started in the framework of the EU H2020 R&DI programme. In: Proceedings of the 33rd CIB W78 Conference 2016, Oct. 31st – Nov. 2nd 2016, Brisbane (2016)

Mohammadi, M.: Bayesian evaluation and comparison of ontology alignment systems. IEEE Access **7**, 55035–55049 (2019)

Mohammadi, M., Atashin, A.A., Hofman, W., Tan, Y.: Comparison of ontology alignment systems across single matching task via the McNemar's test. ACM Trans. Knowl. Discovery from Data (TKDD) **12**(4), 1–18 (2018a)

Mohammadi, M., Hofman, W., Tan, Y.H.: A comparative study of ontology matching systems via inferential statistics. IEEE Trans. Knowl. Data Eng. **31**(4), 615–628 (2018b)

Mohammadi, M., Rezaei, J.: Evaluating and comparing ontology alignment systems: an MCDM approach. J. Web Semantics **64**, 100592 (2020)

Nguyen, T.T.A., Conrad, S.: Ontology matching using multiple similarity measures. In: 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K), vol. 1, pp. 603–611. IEEE, Piscataway (2015)

Niu, Z., Yu, Z., Tang, W., Wu, Q., Reformat, M.: Wind power forecasting using attention-based gated recurrent unit network. Energy **196**, 117081 (2020)

Olaofe, Z.O., Folly, K.A.: Wind energy analysis based on turbine and developed site power curves: a case-study of darling city. Renew. Energy **53**, 306–318 (2013)

Ouali, I., Ghozzi, F., Taktak, R., Sassi, M.S.H.: Ontology alignment using stable matching. Procedia Comput. Sci. **159**, 746–755 (2019)

Pauwels, P., Terkaj, W.: EXPRESS to OWL for construction industry: towards a recommendable and usable ifcOWL ontology. Automat. Construct. **63**, 100–133 (2016). https://doi.org/10.1016/j.autcon.2015.12.003. https://www.sciencedirect.com/science/article/pii/S0926580515002435

Qdr, Q.: Benefits of demand response in electricity markets and recommendations for achieving them. US Department, Energy, Washington, Technical Report (2006)

Seeger, M., Salinas, D., Flunkert, V.: Bayesian intermittent demand forecasting for large inventories. In: Proceedings of the 30th International Conference on Neural Information Processing Systems, pp. 4653–4661 (2016)

Shojaeighadikolaei, A., Ghasemi, A., Jones, K.R., Bardas, A.G., Hashemi, M., Ahmadi, R.: Demand responsive dynamic pricing framework for prosumer dominated microgrids using multiagent reinforcement learning. In: 2020 52nd North American Power Symposium (NAPS), pp. 1–6. IEEE, Piscataway (2021)

Siano, P.: Demand response and smart grids—a survey. Renew. Sust. Energ. Rev. **30**, 461–478 (2014). https://doi.org/10.1016/j.rser.2013.10.022. https://www.sciencedirect.com/science/article/pii/S1364032113007211

Stap, R., Daniele, L.: Ontology 'a CIM profile for smart grids'. https://ontology.tno.nl/cerise/cim-profile/. Accessed 07 Feb 2021

Sun, Z., Sun, H., Zhang, J.: Multistep wind speed and wind power prediction based on a predictive deep belief network and an optimized random forest. Math. Prob. Eng. **2018** (2018)

Valdes, J., Poque González, A.B., Ramirez Camargo, L., Valin Fenández, M., Masip Macia, Y., Dorner, W.: Industry, flexibility, and demand response: applying German energy transition lessons in Chile. Energy Res. Soc. Sci. **54**, 12–25 (2019). https://doi.org/10.1016/j.erss.2019.03.003. https://www.sciencedirect.com/science/article/pii/S2214629618306054

Wang, H., Lei, Z., Zhang, X., Zhou, B., Peng, J.: A review of deep learning for renewable energy forecasting. Energy Convers. Manag. **198**, 111799 (2019)

Wicaksono, H.: An integrated method for information and communication technology (ICT) supported energy efficiency evaluation and optimization in manufacturing: knowledge-based approach and energy performance indicators (EnPI) to support evaluation and optimization of energy efficiency. Ph.D. Thesis, Karlsruher Institut für Technologie (KIT) (2016). https://doi.org/10.5445/IR/1000060882

Wicaksono, H., Dobreva, P., Häfner, P., Rogalski, S.: Methodology to develop ontological building information model for energy management system in building operational phase. In: International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management, pp. 168–181. Springer, Berlin (2013)

Wicaksono, H., Jost, F., Rogalski, S., Ovtcharova, J.: Energy efficiency evaluation in manufacturing through an ontology-represented knowledge base. Intell. Syst. Account. Financ. Manag. **21**(1), 59–69 (2014)

Wicaksono, H., Rogalski, S., Kusnady, E.: Knowledge-based intelligent energy management using building automation system. In: 2010 Conference Proceedings IPEC, pp. 1140–1145. IEEE, Piscataway (2010)

Xie, C., Chekol, M.W., Spahiu, B., Cai, H.: Leveraging structural information in ontology matching. In: 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), pp. 1108–1115. IEEE, Piscataway (2016)

Yassine, A.: Implementation challenges of automatic demand response for households in smart grids. In: 2016 3rd International Conference on Renewable Energies for Developing Countries (REDEC), pp. 1–6. IEEE, Piscataway (2016)

Zhou, L., Cheatham, M., Krisnadhi, A., Hitzler, P.: A complex alignment benchmark: GeoLink dataset. In: International Semantic Web Conference, pp. 273–288. Springer, Berlin (2018)

The Influence of Cognitive Biases in Production Logistics

Florian Knapp, Melanie Kessler, and Julia C. Arlinghaus

Abstract Digitalization and the transformation of industry into Industry 4.0 is changing the character of production logistics substantially. New Logistics 4.0 technologies are largely enabling automated decision-making by machines. Human decisions are nevertheless still required. Research shows, however, that human decisions are often more biased and less rational than most logistics models assume. Decision makers and decision support system designers therefore need to understand the influence of the so-called cognitive biases on the human decision-making process. We contribute to the scholarship on this issue by combining the literature streams of logistics and cognitive biases. We demonstrate the influence of cognitive biases on human decision-making based on typical decisions in logistics and derive initial hypotheses.

1 Introduction

Industry 4.0 or the Fourth Industrial Revolution is changing all areas of manufacturing. New technologies, such as online cyber-physical systems, the Internet of Things (IoT), and cloud-based solutions, are transforming conventional man-

---

Julia C. Arlinghaus was affiliated to LogDynamics until 2017 as Professor of Optimization of Production and Logistics Networks at the Jacobs University Bremen, Germany

F. Knapp
Otto von Guericke University, Magdeburg, Germany
e-mail: florian.knapp@ovgu.de

M. Kessler
RWTH Aachen University, Aachen, Germany
e-mail: melanie.kessler@rwth-aachen.de

J. C. Arlinghaus (✉)
Otto von Guericke University, Magdeburg, Germany
e-mail: julia.arlinghaus@ovgu.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_9

ufacturing facilities into the so-called smart factories (Skapinyecz et al. 2018). These technologies are enabling industry to transition to full digitalization and smart manufacturing processes (Erol et al. 2016; Brettel et al. 2014; De Felice and Petrillo 2012). As digitalization increases, the field of logistics is also undergoing a major transformation (Rushton et al. 2006). Logistics is a broad field comprising the integrated planning, implementation, and control of the flow and storage of materials and products, services, information, energy, and other resources. This means that every flow into, through, and out of an organization is covered from its point of origin to its point of consumption with the aim of meeting customer demands (Johnson et al. 1999). The transformation of traditional logistics in conjunction with Industry 4.0 generally referred to as Logistics 4.0 offers logisticians new opportunities to boost efficiency and cut costs and creates new opportunities through the aforementioned digital innovations (Bamberger et al. 2017).

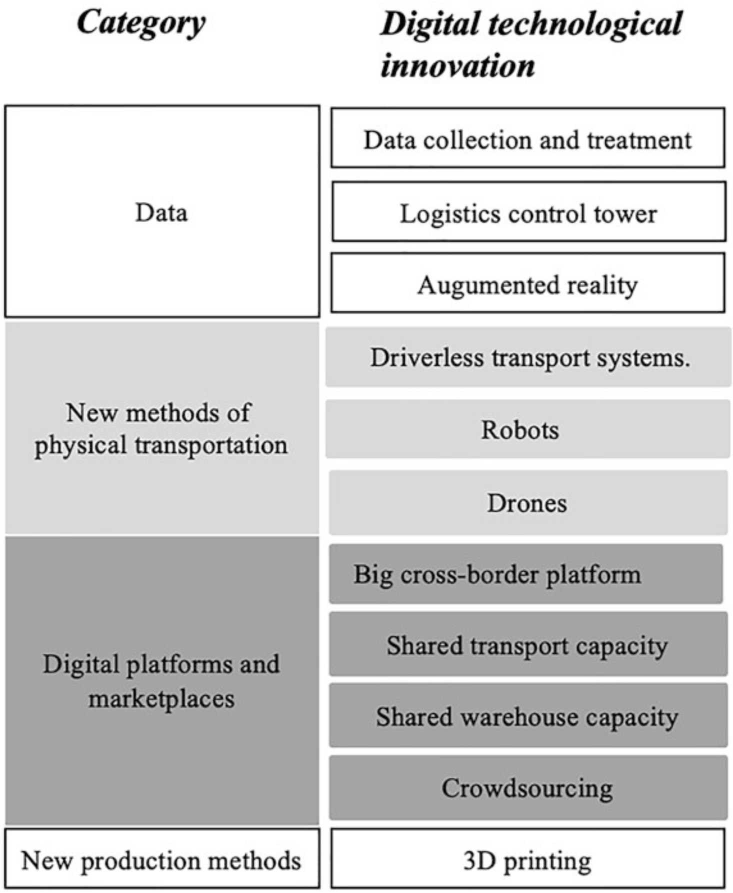
Figure 1 summarizes the new domains based on Logistics 4.0 and the new opportunities these new technologies provide. No longer confined to one company, industry or country, material, information and services will increasingly become a global system, as the lower half of Fig. 1 indicates. The range of activities constitutes a challenge for all businesses, though.

Logistics in general chiefly focuses on the optimization of business activities, i.e., improving the flow and storage of inventory, goods, and services through the supply chain (CSCMP 2021). Improving logistics performance in relation to the aforementioned parameters requires consideration of multiple factors.

The new technologies presented in Fig. 1 are intended to enable these logistical parameters' complex connections and assist humans with their decision-making in various logistics operations. Whereas logistics research has substantial experience with these standard rational factors, it frequently overlooks human factors. Judgment and decision-making are crucial and fundamental to the field of logistics, but humans tend to make systematic errors when making decisions, especially when they are encumbered by time pressure and uncertainty (Bazerman et al. 2002; Stanovich and West 1998). Cognitive biases are the root cause of this (Tversky and Kahneman 1974a). The phenomenon of lead time syndrome (LTS), for instance, illustrates how biased human decision-making diminishes logistics performance. Bendul and Knollmann (2016) identified several cognitive biases that influence production planners in the event they are confronted by unforeseen events, such as declining due date reliability. Planners consequently tend to adjust system parameters, such as planned lead times, in order to improve the logistics performance, albeit this actually worsens due date reliability.

This example demonstrates the effect of cognitive biases on logistics performance and illustrates the strong interrelationship between logistics and manufacturing operations. We therefore focus on production logistics in this book chapter with the intention of answering the following research question: *Which types of decisions in production logistics are influenced by cognitive biases and how do these biases affect production logistics performance*?

To this end, we combine the literature streams on production logistics and cognitive biases. Considering practical examples, we map the cognitive biases identified beforehand to decision situations typical to production logistics and



Category	Digital technological innovation
Data	Data collection and treatment Logistics control tower Augumented reality
New methods of physical transportation	Driverless transport systems. Robots Drones
Digital platforms and marketplaces	Big cross-border platform Shared transport capacity Shared warehouse capacity Crowdsourcing
New production methods	3D printing

Fig. 1 Logistics 4.0 building blocks (Skapinyecz et al. 2018)

demonstrate how they influence human decision-making. After examining the effects of cognitive biases, we derive initial propositions for further research in this field.

2 Literature Review

2.1 The Human Decision-Making Process and Cognitive Biases

The human decision-making process has been the focus of research in various fields, such as psychology and strategic and behavioral organization management. A

Table 1 Relevant cognitive biases

Category of cognitive bias	Description
Memory bias	*Group of cognitive biases related to information storage and availability*
Statistical bias	*Tendency to overestimate or underestimate certain statistical parameters*
Adjustment bias	*Tendency to stick to the first information available or to a reference point*
Presentation bias	*Decision based on the display of information*
Situational bias	*The way a person responds to a general situation*

decision is a situational response that consists of three parts (Tversky and Kahneman 1974a):

1. More than one possible choice of action is under consideration in the choice set.
2. The decision maker can form expectations of future events and outcomes ensuing from each course of action, which are describable by degrees of belief or probabilities.
3. The consequences of the possible outcomes are assessable on an evaluative continuum determined by current goals and personal values.

The broad scope and various linked variables can make decisions in production logistics very complex. Most decisions have to be made to optimize business activities and, in some cases, only have an impact after a prolonged period. Tversky and Kahneman (1974b) demonstrated that humans make systematic errors in particular decision environments and introduced the term *cognitive bias* to denote subconscious errors in human decision-making. Reflecting on cognitive biases and their impact is essential during the decision-making process. Individuals make irrational decisions based on their respective backgrounds, knowledge, experiences, and attitudes (Bendul and Zahner 2019; Arlinghaus et al. 2020). We mapped several cognitive biases from each category of the most important main groups of Arnott's (MacCarthy and Fernandes 2000) categorization, some of which examined in more detail in this study, to the three decision types in production logistics. Table 1 shows the categorization of cognitive biases effects by Arnott (MacCarthy and Fernandes 2000) and in which decision-making area they have an impact.

2.2 *Decisions in Production Logistics*

Production logistics describes planning, coordination, transportation, and similar activities related to manufacturing. Production logistics does not function in any set way since it is contingent on each company's manufacturing environment and products (Jonsson and Mattson 2003; Fleischmann et al. 2005). This concurs with

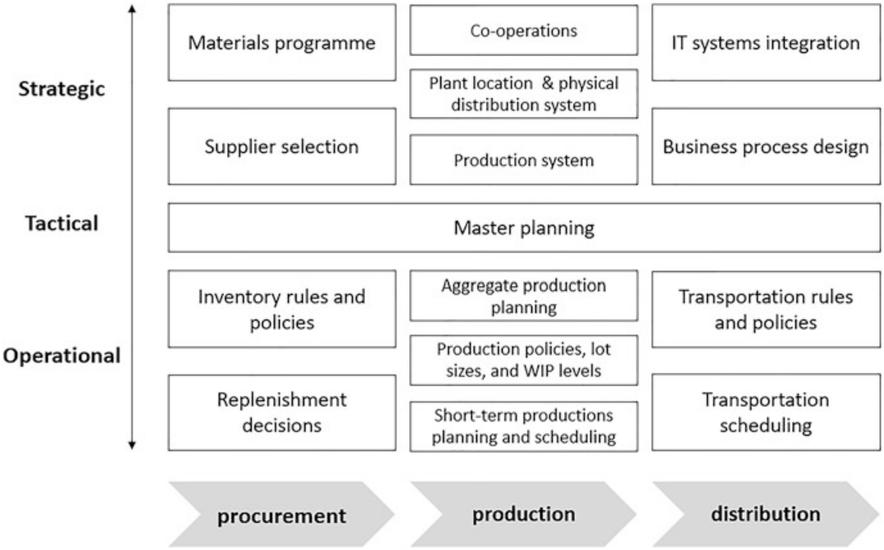


Fig. 2 Production logistics decision areas (based on Fleischmann et al. (Arnott 2006))

Semini et al. (2006) who point out that “manufacturing logistics encompasses aspects of several overlapping fields, including operations and production management, logistics and supply chain management, and advanced planning.” This field’s diversity has spawned an abundance of different approaches. Some authors, such as Chan (2005), view production planning and control as the crux of production logistics. Others, such as Strandhagen et al. (2006), employ a modified version of the supply chain planning matrix developed by Fleischmann et al. (Arnott 2006).

Figure 2 presents the range of production logistics that require human decision-making. Decision types can be categorized as strategic, tactical, and operational based on the decision horizon. Strategic decisions typically have a long-term horizon, operational decisions a short-term horizon, and tactical decisions a medium-term horizon.

We added the decision horizon to the supply chain planning matrix of Fleischmann et al. (Arnott 2006) and mapped the different decision types to the typical tasks in production logistics as illustrated in Fig. 2.

Production logistics decisions are classifiable by the decision horizon introduced (Ghiani et al. 2004). Decisions are made iteratively and hierarchically from the strategic to the tactical and the operational level. These decision-making levels are identified and explained with examples in Table 2 (Seifi 2011).

Table 2 Decision-making levels with time horizon and level attributes

Decision-making level	Time horizon	Decision maker	Decisions	Data density and basis
Strategic	*> 1 year*	*Executive management*	*Capital reduction*	*Imprecise*
		Top management	*Cost reduction*	*Incomplete*
		Stockholders	*Service-level improvement*	*Requires forecasts*
Tactical	*Monthly*	*Middle management*	*Production planning*	*Disaggregated*
	Quarterly	*Logistics engineers*	*Resource planning*	
	Annually		*Transportation planning*	
Operational	*Daily*	*Supervisors*	*Vehicle loading*	*Large quantities of data*
	Weekly		*Dispatching*	
			Shipping	
			Warehouse routines	

3 Conceptual Framework of Distorted Human Decision-Making in Production Logistics

3.1 Strategic Decisions

Memory Biases: Imaginability Bias *Imaginability bias* describes individuals' assumption that events they can easily imagine are more probable. Individuals' own attitudes and imagination thus influence the assessment of the probability of a risk occurrence (Taylor and Thompson 1982), A company is more likely to install an integrated IT system to track spare parts across units if the decision maker is IT savvy than if the decision maker is incapable of imagining such a system or any added value from its implementation.

Adjustment Biases: Conservatism Bias Individuals that weight new information less than initial information are displaying *conservatism bias* (Pompian 2012). In the case of strategic decision this often occurs that persons stick to already known and common solutions and avoid completely new methods. For instance, they might pay less attention to new production logistics processes such as additive manufacturing despite all its advantages.

Confidence Biases: Confirmation Bias The reflection of a decision maker's attitude in a decision stems from *confirmation bias*. Arguments that support personal opinion are weighted more heavily than others that do not (Wheeler and Arunachalam 2008). For instance, buyers might select vendors based on their personal preferences (Table 3).

Since some of the many decisions made at the strategic level are fraught with uncertainty, they depend on the decision maker's skills and background. We therefore propose:

P1: Cognitive biases related to the decision maker's skills and background heavily influence strategic decisions based on uncertain data.

Table 3 Cognitive biases in strategic decisions

	Main decision(s) at the strategic level
Cognitive biases	Which relevant facts deserve further consideration?
Memory bias	
• Imaginability bias	*Am I the best person to judge?*
Adjustment bias	
• Conservatism bias	*What changed since the last period under review?*
Confidence bias	
• Confirmation bias	*Are there other targeted options?*

3.2 Tactical Decisions

Presentation Biases: Ambiguity Effect The *ambiguity effect* describes humans' tendency to favor the seemingly simplest option over more complex options (Ellsberg 1961). Practitioners prefer quick and easy solutions that can be found and applied quickly. Exemplarily, the lead time syndrome shows this connection. When facing decreasing due date reliability, the most simple looking option is to adapt planned lead times. Other influencing effects are consequently underestimated.

Confidence Biases: Illusion of Control The *illusion of control* describes individuals' tendency to overestimate their ability to solve difficult problems (Brenner et al. 1996). People tend to overestimate their own abilities as well as their own plans. This is especially true when production and logistics parameters are monitored insufficiently because the decision maker deems them unimportant.

Situational Biases: Ostrich Effect The *ostrich effect* describes people's habit of ignoring obviously negative information in order to advance certain (e.g., their own) interests (Karlsson et al. 2009). For instance, the workload of a machine that is heavily utilized but also prone to malfunction is not reduced and the risk of jeopardizing due date reliability is ignored (Table 4).

Although the base data can be used for decisions, the decision maker's (personal) skills also influence decision depending on its scope. We therefore propose:

P2: Both cognitive biases that are related to personality factors and influence the treatment of data influence medium-term decisions.

3.3 Operational Decisions

Adjustment Biases: Anchoring Effect This cognitive bias describes the treatment of initial information as an "anchor" to which individuals hold fast (Tversky and Kahnemann 1974). This might result in new technologies, such as automated guided vehicles, not being considered in transportation planning, for instance.

Table 4 Cognitive biases in tactical decisions

	Main decision(s) at the tactical level
Cognitive biases	Which relevant facts deserve further consideration?
Presentation bias	
• Ambiguity effect	*Are there other options that will create future added value?*
Confidence bias	
• Illusion of control	*Is the project, schedule or the like still realistic?*
Situational bias	
• Ostrich effect	*Is all information assessed equally?*

Table 5 Cognitive biases in operational decisions

	Main decision(s) at the operational level
Cognitive biases	Which relevant facts deserve further consideration?
Adjustment bias	
• Anchoring effect	*What changed during a period of time?*
Statistical bias	
• Correlation bias	*Has the correct context been factored in everywhere?*
Situational bias	
• Complexity effect	*Is there useless data?*

Statistical Biases: Correlation Bias Humans assessing two concurrent risks tend to overestimate their probability of occurrence when they have occurred previously (Kahneman 2002). This is called *correlation bias* and makes it difficult to identify triggers since a correlation of both risks is assumed. Process difficulties or improvements (Event 1) are thus associated with the new solution that has just been implemented (Event 2). Since this generates false assumptions about correlation, it can result in incorrect assumptions, especially in the case of planning updated daily because individuals might believe that one parameter indicates good schedule adherence for a product, even though it has no real correlation to schedule adherence.

Situational Biases: Complexity Effect The *complexity effect* describes individuals' bias when they are under time pressure or overloaded by information (Ordonez and Benson 1997). Their assessment of various factors' effects and potential impact is consequently unduly complex since they often link various parameters. This can result in incorrect decisions, especially in conjunction with time pressure, e.g., in replenishment because of shortages or unduly high inventory levels, and thus in high costs for storage space or problems caused by missing material (Table 5).

Since the decision maker can access a large quantity of data at the operational level, we propose that:

P3: As data density increases, these cognitive biases affect decisions about uncertainty, data handling, and time pressure adversely.

4 Conclusion

We identified potential decisions in production logistics and correlated them with potential cognitive biases in order to answer to the research question introduced: *Which types of decisions in production logistics are influenced by cognitive biases and how do these biases affect production logistics performance*? After identifying a correlation between the decision-making level and the possible type of bias, we advanced three propositions that provide initial insights on potential correlations in decision-making in production logistics. A closer examination of the different levels of production logistics reveals that the main groups of cognitive biases cannot be assigned to specific levels in the decision-making process. Categorizing cognitive biases based on the density of data for the potential decision would constitute a better approach since some effects stem more from data handling and others more from the decision maker's personality.

Since biases correlate with multiple levels of logistics decisions, we see a need for further research in this area. Experiments with practitioners would help assign cognitive biases to levels and decisions better and enable drawing conclusions about the extent to which personal or individual factors influence decisions when all subjects have similar backgrounds and perform identical tasks.

References

Arlinghaus, J.C., et al.: The influence of cognitive biases on supply chain risk management in the context of digitalization projects. In: Freitag, M., Haasis, H.D., Kotzab, H., Pannek, J. (eds.) Dynamics in Logistics. LDIC 2020. Lecture Notes in Logistics, pp. 137–147. Springer, Cham (2020)

Arnott, D.: Cognitive biases and decision support systems development: a design science approach. Inf. Syst. J. **16**(1), 55–78 (2006)

Bamberger, V., et al.: Logistics 4.0 – facing digitalization – driven disruption. PRism. **1**, 38–51 (2017)

Bazerman, M.H., et al.: Cognitions and behavior in asymmetric social dilemmas: a comparison of two cultures. J. Appl. Psychol. **87**(1), 87–95 (2002)

Bendul, J., Knollmann, M.: The human factor in production planning and control: considering human needs in computer aided decision support systems. Int. J. Manuf. Technol. Manag. **30**(5), 346–368 (2016)

Bendul, J.C., Zahner, J.H.: Understanding the influence of cognitive biases in production panning and control. In: Ahram, T.Z., et al. (eds.) Human Systems Engineering and Design II. IHSED 2019. Advances in Intelligent Systems and Computing, vol. 1026, pp. 280–285. Springer, Cham (2019)

Brenner, L.A., et al.: On the evaluation of one-sided evidence. J. Behav. Decis. Mak. **9**(1), 59–70 (1996)

Brettel, M., et al.: Fostering creativity in new product development through entrepreneurial decision making. Creat. Innov. Manag. **23**(4), 351–511 (2014)

Chan, J.W.K.: Competitive strategies and manufacturing logistics – an empirical study of Hong Kong manufacturers. Int. J. Phys. Distrib. Logist. Manag. **35**(1), 20–43 (2005)

CSCMP Supply Chain Management Definition and Glossary. https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx. Last accessed 06 Aug 2021

De Felice, F., Petrillo. A.: Green policy in a manufacturing system. In: IEEE (eds.) 2nd International Conference on Communications, Computing and Control Applications, CCCA, pp. 1–6 (2012)

Ellsberg, D.: Risk, ambiguity, and the savage axioms. Q. J. Econ. **75**(4), 643–669 (1961)

Erol, S., et al.: Tangible industry 4.0: a scenario-based approach to learning for the future of production. Procedia CIRP. **54**, 13–18 (2016) 6th CIRP Conference on Learning Factories

Fleischmann, B., et al.: Supply Chain Management and Advanced Planning, 3rd edn. Springer, Berlin (2005)

Ghiani, G., et al.: Introduction to Logistic Systems Planning and Control. Wiley, West Sussex (2004)

Johnson, D.F., et al.: Contemporary Logistics. Prentice Hall, New Jersey (1999)

Jonsson, P., Mattson, S.A.: The implications of fit between planning environments and manufacturing planning and control methods. Int. J. Oper. Prod. Manag. **23**(8), 872–900 (2003)

Kahneman, D.: A perspective on intuitive judgment and choice: mapping bounded rationality. Novel price literature. Am. Psychol. **58**(9), 697–721 (2002)

Karlsson, N., et al.: The ostrich effect: selective attention to information. J. Risk Uncertain. **38**, 95–115 (2009)

MacCarthy, B.L., Fernandes, F.C.F.: A multi-dimensional classification of production systems for the design and selection of production planning and control systems. Prod. Plan. Control. **11**(5), 481–496 (2000)

Ordonez, L., Benson, L.: Decisions under time pressure: how time constraint affects risky decision making. Organ. Behav. Hum. Decis. Process. **71**(2), 121–140 (1997)

Pompian, M.M.: Conservatism bias. In: Behavioral Finance and Wealth Management: how to Build Investment Strategies that Account for Investor Biases, 2nd edn, pp. 63–71. Wiley, Hoboken, NJ (2012)

Rushton, A., et al.: The Handbook of Logistics and Distribution Management. Kogan Page, London (2006)

Seifi, M.S.: Logistics strategic decisions. In: Logistics Operations and Management, Concepts and Models, pp. 43–53. Elsevier, Amsterdam (2011)

Semini, M., et al.: Applications of discrete-event simulation to support manufacturing logistics decision-making: a survey. In: Perrone, L.F., et al. (eds.) Proceedings of the 2006 Winter Simulation Conference, pp. 1946–1953. WSC, Monterey, CA (2006)

Skapinyecz, R., et al.: Logistic aspects of industry 4.0. IOP Conf. Ser. Mater. Sci. Eng. **448**, 012014 (2018)

Stanovich, K.E., West, R.E.: Individual differences in rational thought. J. Exp. Psychol. Gen. **127**(2), 161–188 (1998)

Strandhagen, J. O. et al.: Which manufacturing logistics decisions are supported by operational research? A literature survey. IFIP International Federation for Information Processing (2006)

Taylor, S.E., Thompson, S.C.: Stalking the elusive "vividness" effect. Psychol. Rev. **89**(2), 155–181 (1982)

Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science. **18**(4157), 1124–1131 (1974a)

Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science. **18**(4157), 1112–1131 (1974b)

Tversky, A., Kahnemann, D.: Judgement under uncertainty: heuristics and biases. Science. **185–415**, 1124–1131 (1974)

Wheeler, P.R., Arunachalam, V.: The effects of decision aid design on the information search strategies and confirmation bias of tax professionals. Behav. Res. Account. **20**(1), 131–145 (2008)

Part III

Fields of Application in Logistics

Automobile Logistics 4.0: Advances Through Digitalization

Marit Hoff-Hoffmeyer-Zlotnik, Michael Teucke, Stephan Oelker, and Michael Freitag

Abstract In today's buyer's markets, logistical service quality is of great importance, particularly for high-priced products. For many buyers, the automobile is the epitome of a high-priced, individually customized product. Hence, customers expect a high delivery service quality. Distribution logistics provides the link between the manufacturer and the customer and is therefore responsible for providing these services. In automobile logistics, the distribution chain involves several stakeholders, such as manufacturers, transport, and technical services providers, as well as dealers. Seamless coordination between them is crucial. Digitalization is an inevitable means for gaining a high level of transparency among the stakeholders. In addition, it provides technologies for developing effective and tailor-made assistance and control systems that support processes along the distribution chain. In recent years, the BIBA—Bremer Institut für Produktion und Logistik GmbH has been active in automobile logistics research and has been involved in developing several assistant and control systems that range from relatively simple track and trace systems to highly complex compound control systems. In this paper, we present an overview of the developed systems and provide a vision, how recent and ongoing research improves automobile logistics and how to further potential can be leveraged.

M. Hoff-Hoffmeyer-Zlotnik (✉) · M. Teucke · S. Oelker
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany
e-mail: hhz@biba.uni-bremen.de; tck@biba.uni-bremen.de; oel@biba.uni-bremen.de

M. Freitag
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

Faculty of Production Engineering, University of Bremen, Bremen, Germany
e-mail: fre@biba.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_10

1 Introduction and Motivation

1.1 Research Motivation

Automobiles constitute high-priced, custom-built consumer goods. Customers expect a high delivery service quality when buying a new car (Klug 2018). The delivery service is an aggregate measure of delivery time, delivery reliability, delivery quality, and delivery flexibility (Pfohl 2010). These features influence customer satisfaction and are thus decisive for the economic success of companies (Seeck 2010). For car manufacturers, it is crucial to deliver vehicles to the customers on time while also meeting quality requirements (Klug 2018; Ambe and Badenhorst-Weiss 2010).

Ensuring a high delivery service quality is the task of distribution logistics. As the link between production and the customer, distribution is responsible for satisfying customer needs by providing customers with the goods they want (Ross 2015).

Overcoming the challenges listed above is particularly difficult when focal companies dominate the order fulfillment process, as it is the case with the car manufacturers in the automotive industry (Scholz-Reiter et al. 2008). Car manufacturers have primarily pursued the goal of optimizing their production processes, whereas the holistic optimization of the whole value chain has been neglected (Hermes 2011; Scholz-Reiter et al. 2008).

As a result, data for distribution planning is often transmitted to the logistics service providers (LSP) incompletely or late. In addition, the actual number of vehicles to be transmitted for transport on the individual routes often deviates significantly from the planned numbers (Holweg and Miemczyk 2002). Therefore, it is not possible to start planning and controlling the activities required for order processing at a sufficiently early stage. Transport planning can usually only be carried out after the end-of-line area when the vehicles are handed over to vehicle logistics. As a result, optimization of the transports is only insufficiently possible.

A further complication exists if the logistics service providers commissioned with distribution are inadequately connected to the manufacturers' information technology (IT) systems (Holweg and Miemczyk 2002). Many vehicle distribution processes are characterized by isolated heterogeneous IT systems, where the companies involved store their data and only forward it in parallel with the material flow (Ruthenbeck et al. 2010). In addition, media discontinuities often occur in the distribution network due to outdated IT systems and the lack of compliance with standards, which generates additional efforts for data entry and thus sources of error (Ruthenbeck et al. 2010). The externalization of logistics services (Gudehus 2010; Klug 2018; Schuh 2013) increases the need for better IT connection and efficient interfaces between the companies involved in distribution (Scholz-Reiter et al. 2009b).

Increasing digitalization in the context of Industry 4.0 offers the potential to improve transparency in supply chains through real-time or near real-time data (Musa et al. 2014; Genc et al. 2014). Industry 4.0 aims to integrate cyber-physical systems (CPS) in conjunction with the application of the Internet of Things in production and logistics in order to increase competitiveness (Kagermann et al. 2013). Cyber-physical systems, the Internet of Things, and Digital Twins (Negri et al. 2017) can achieve data availability and data quality to ensure efficient supply chain management (Monostori 2018). The following sub-section provides an overview of the research contribution and the structure of the chapter.

1.2 Research Contribution

In recent years, the BIBA—Bremer Institut für Produktion und Logistik GmbH, a member of the Bremen Research Cluster for Dynamics in Logistics (LogDynamics 2021), has done extensive research on increasing information transparency at transshipment points and across the distribution chain, using technologies related to Industry 4.0. In addition, the research includes the development of assistance and control systems that use similar technologies and improve planning and control along the vehicle distribution chain. BIBA's research work addressing the challenges of vehicle distribution was application-driven, meaning it was carried out from the perspective of how Industry 4.0 related technologies and systems can be applied to the planning, control, and execution of logistic processes. The purpose of this chapter is to provide a summary and an overview of this work, its results achieved so far, and the resulting benefits for vehicle distribution logistics.

This chapter is structured as follows: This first section provides an introduction to the topic. The second section describes the processes of finished vehicle logistics and the related planning and control. The third section describes the available digital technologies that can support transparency in distribution processes as well as their planning and control. The fourth section describes several selected assistance and control systems developed by or with the support of BIBA or within LogDynamics and used digital technologies to improve the transparency of vehicle distribution processes. The fifth section points out potential benefits and improvements in vehicle logistics that result from these applications. The final section provides a summary and an outlook on ongoing and future research activities in finished vehicle logistics.

2 Finished Vehicle Logistics

2.1 Tasks of Distribution Logistics

Distribution logistics provides the link between the manufacturer and the customer (Fig. 1). Functions include planning, managing, executing, and controlling material and information flows between the end of production and handover to the customer (Pfohl 2010).

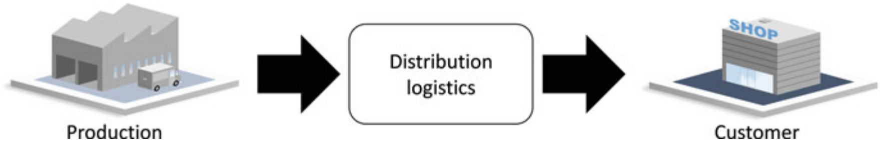


Fig. 1 Distribution logistics (adapted from Filz 1989)

Distribution logistics is divided into order processing, warehousing, and transport (Filz 1989). From the customer's point of view, distribution logistics is vital, as it represents the actual service provision (Hammer and Champy 2003). Within the distribution of finished vehicles, special features are, on the one hand, that specialized load carriers are used and, on the other hand, that the transported goods themselves can move. Therefore, the transport is called roll-on-roll-off (RoRo) transport (Schukraft et al. 2018).

In the automotive sector, distribution logistics begins at the original equipment manufacturer's (OEM) site after the vehicle's production has been completed and the vehicle has successfully passed the quality inspection (Scholz-Reiter et al. 2012b). There are two alternatives to how the product can reach the customer. On the one hand, in relatively few cases, the finished vehicles are collected from the factory by the customers. This case is not considered further in this article. On the other hand, the finished vehicles are prepared for transport according to their destination and transported to a dealer via a multi-stage multi-modal transport chain. This process makes up a significant fraction of the time from production to reaching the owner; just the crossing from Europe to the USA takes about two weeks, while the production takes only about one day. The distribution process involves the OEM, transport, and technical service providers, and the dealer as stakeholders (Gudehus 2012). The transport process must run as planned because a vehicle is a highly individualized product and is usually customized. Accordingly, another product cannot easily substitute it if it does not arrive at the customer as planned.

Starting from the OEM's site, logistics service providers transport the vehicles from one transshipment point to the next. The logistics service providers have special modes of transport at their disposal and differentiate between surface transport (truck) and en-route transport (especially rail and ship). In long-distance traffic, the transports for several OEMs are usually combined so that vehicles from different manufacturers can be found on the same ship.

Figure 2 illustrates an overview of the distribution network in Europe. An exemplary distribution chain for finished vehicles is a three-stage transport chain. The finished vehicles leave the OEM and are transported from there by train or truck to a transshipment point (pre-carriage). From there, a ship, train, or truck transports the vehicles to a second transshipment point, which is as close as possible to the destination of the vehicles (main-carriage). Afterward, the vehicles are usually transported by truck to the dealer (on-carriage). At each transshipment point, a quality check is carried out. In general, international vehicle distribution mainly

Fig. 2 Exemplary distribution chains for finished vehicle distribution (adapted from Haubrich 2017)

takes place by ship. Within Europe, rail is mainly used, as it is cheaper than a truck (Klug 2018; Herold 2005).

The transshipment points for finished vehicles play a special role in the finished vehicle supply chain. The so-called vehicle compounds can handle several million vehicles per year, load and unload over one thousand ships per year and store up to 100,000 vehicles on site (Klug 2018). In addition, to storing and turning over vehicles, technical services are often offered at the transshipment points. These services, which can be booked at short notice, include, for example, washing and removal of the transport protection film, installation of special equipment and accessories, or technical modifications (Fischer 2004). This means that customer wishes can be implemented relatively late in the process chain. In addition, slight damage to the vehicles can be repaired immediately, and the vehicles do not have to be transported to a repair shop or back to the OEM.

Dealers are the last point in the supply chain before the vehicle is handed over to the customer. The delivery to the dealers is usually made by truck from the last transshipment point (Ruthenbeck et al. 2010). The dealers are in direct contact with the end customer, so they need to know when the vehicles will arrive at their premises. They can inform the customers about the progress and arrange a handover date. The next sub-section describes the processes and challenges for planning and controlling the logistics of finished vehicles.

2.2 Planning and Control Within Finished Vehicle Logistics

Planning and control consider the temporal, quantitative, and, if necessary, spatial aspects of production and logistics processes. Planning makes decisions on anticipated future operations to prepare their best possible execution. For instance, planning identifies alternative process executions, evaluates the alternative options, and selects the best available option. Strategic, tactical, and operational planning correspond to a long, medium, and short-term time planning horizon. Control takes place in parallel with the processes and allows for intervention in case deviations from the plan have occurred in the meantime (Lödding 2011).

Within the distribution of finished vehicles, many different players are involved, and therefore, the planning and control of the entire distribution chain pose a challenge. In terms of day-to-day coordination among the different companies, operational planning and control levels are of particular interest. The process starts at the manufacturers. Based on order information from the dealers, they plan the distribution routes and monitor and coordinate the distribution along the entire distribution chain (Klug 2018). As soon as the vehicles enter the delivery process, the manufacturer is in contact with the logistics service providers to exchange information about transport, service, blockings, confirmations, entry and exit bookings, loss, damage, rework, and technical services (Schenk and Clausen 2020). However, a uniform IT system for standardized information exchange across all companies does not exist (Schenk and Clausen 2020). All parties work with their systems and with different levels of formality, automation, and assistance.

Vehicle compounds are of particular importance within distribution logistics as they serve as intermediate storage facilities, deconsolidation, and consolidation points and offer technical services (Klug 2018). Operational planning and control of a vehicle compound involve, among others, the following tasks: (i) allocation of loading and unloading points to specific modes of transport, (ii) parking space allocation for the vehicles with the aim of short travel distances on the compound, (iii) planning and control of the technical service centers, and (iv) assignment of specific transport orders to the driving personnel (Hoff-Hoffmeyer-Zlotnik et al. 2017; Görges and Freitag 2019b). However, the information that is available to the vehicle compounds beforehand is limited. For example, detailed information, such as the exact number of vehicles, the required technical services, or the forwarding travel date, might be missing. Thus, detailed planning on the level of single vehicles is challenging, and actions often need to be taken spontaneously and in a flexible manner (Haubrich 2017).

Vehicle compounds are located between the manufacturers and dealers (Fig. 3). While manufacturers serve the distribution chain forecast-driven and via push principles, the dealers' goal is to take the finished vehicles based on demand and thus according to the pull principle (Dias et al. 2010). As a result, the vehicle compounds fulfill a buffer function between the automobile manufacturer and the dealer and thus enable a particularly high degree of flexibility within the distribution chain. However, this comes at the cost of an exceptionally high planning and control

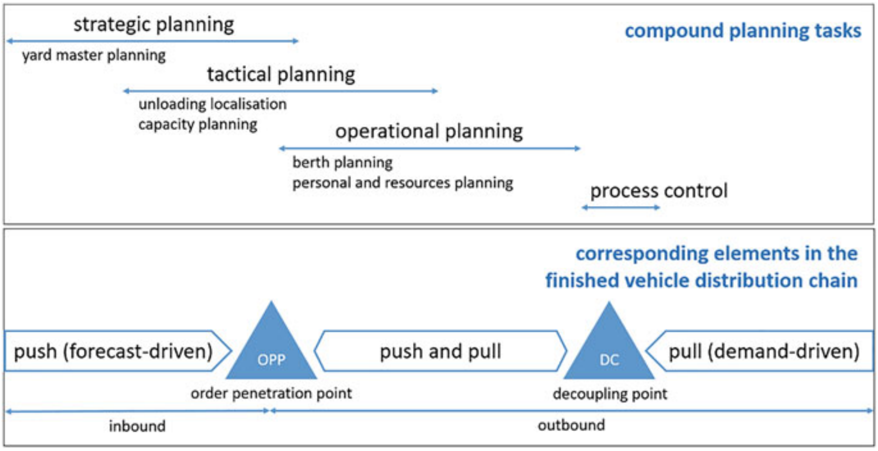


Fig. 3 Planning and control at vehicle compounds (adapted from Görges and Freitag 2019b)

effort for the vehicle compounds (Görges and Freitag 2019a; Görges and Freitag 2019b). Therefore, the next section describes different technologies for providing information and enabling transparency along the distribution chain.

3 Technological Basis for Generating Transparency

In order to obtain transparency about the current status of the process chain, corresponding data is required. This data can, for example, provide valuable information about the current position or condition of the vehicle. From this, it can be derived whether the process is still within the planned state or whether measures need to be taken. Different technologies have been established to collect this data, which will be briefly presented in the following. In addition, the standards for data exchange are also described shortly. An overview of the presented technologies and standards is given in Fig. 4.

3.1 Automatic Identification Technologies

Automatic identification systems are technologies for identification, data acquisition, data collection, and data transmission. The focus is on the assignment of two independent objects. As an example, a number is assigned to a vehicle. Automatic identification is understood to mean the automation of data input into a computer or a communication system. This assignment, i.e., the coding, takes place by means

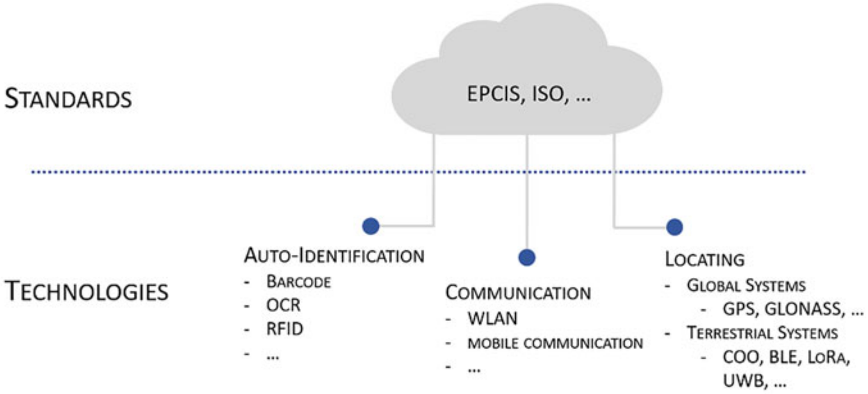


Fig. 4 Overview of technologies and standards that are the basis to gain transparency

of automatic identification. In the following, barcode, OCR, and RFID are briefly introduced, as these technologies are most frequently used in the automotive sector.

3.1.1 Barcode

The barcode is a 1D code, and the functional principle is similar to Morse code. Accordingly, in the simplest form, there are thick bars, thin bars and spaces, which are then interpreted into characters accordingly (Schenk 2021).

The advantages of the barcode are low label/direct marking costs, a high degree of standardization, and a high level of integration into the logistics value chain (Jesse and Rosenbaum 2000). Moreover, metal, water, and heat do not influence it. The disadvantages are that a barcode requires a direct line-of-sight, is susceptible to contamination, has low data capacity and density, and is inflexible regarding the subsequent change (Jesse and Rosenbaum 2000). Nevertheless, they have high acceptance in practice.

3.1.2 Optical Character Recognition

Optical character recognition (OCR) refers to automated text recognition or automatic character recognition within images. In principle, optical recording, for example, delivers a raster graphic as a result with a camera. Text recognition aims to recognize the characters recorded with the image and transfer them into a different format, e.g., plain text, with which further systems can work. For example, in the automotive sector, the vehicle identification number can be photographed and automatically converted into a machine-readable format (Mori et al. 1992).

3.1.3 Radio-Frequency Identification

Radio-frequency identification (RFID) is the term used for transmitter-receiver systems. By transmitting data using radio waves, RFID enables automatic and contactless identification on the one hand and the location of objects on the other. An RFID system always consists of two components:

- A transponder, which is also called a tag. The tag consists of a coupling element, a transmitting/receiving antenna, and a microchip on which data can be stored. In the automotive sector, for example, the vehicle identification number can be stored on the microchip.
- In addition, there is the reader, which can be permanently installed, e.g., in a gate, or mobile, e.g., a handheld reader (Finkenzeller 2015).

RFID systems can be divided into active and passive systems. The main difference is that in an active system, the tags have their own power supply. This usually increases the range and accuracy. However, these systems require more maintenance and are usually more expensive than passive systems. Passive systems do not have their own power supply. The transponder, therefore, draws the energy it needs from a field generated by the reader. The energy transmission and the data transfer take place completely contactless. The passive systems are usually cheaper, but the range and accuracy are generally lower than an active system (Finkenzeller 2015).

3.2 Location (Geopositioning) Technologies

Location finding systems can determine the location of objects in two-dimensional or three-dimensional space. A large variety of systems for automatic location are available, which differ in their geographical scope (or range) and their purpose (Nait-Sidi-Moh et al. 2013).

The location of an object may be computed using different principles. Triangulation, for instance, involves forming triangles to the object from known reference points, establishing a combination of distances and angles. In contrast, multi-lateration determines the distances between an object and several known reference points (Strang et al. 2008).

Most systems use radiolocation based on the known characteristics of radio waves (e.g., field strength or propagation speed) to encode the necessary data. This includes measurement technologies like time-of-flight, round-trip-times, and received signal strength (RSSI) for multi-lateration and angle-of-arrival (AoA) estimation for triangulation. Another measurement technique is fingerprinting, which involves comparing the RSSI or other properties of a signal with the previously mapped local profile (fingerprint) of that same signal. For local or indoor applications, optical technologies (like visible or infra-red light), magnetic fields, or ultrasound may be used besides radiolocation (Strang et al. 2008).

The geographical area for which a system offers its location service differs very strongly between global coverage or extended regions to local coverage of a restricted terrain and a single building.

3.2.1 Global Navigation Satellite Systems

Global navigation systems provide global or regional positioning data using satellites circling around the earth at known altitudes. Thus, they are also known as global navigation satellite systems (GNSS). The satellites transmit their geospatial position as well as time data by radio. Based on the received time signals of at least four satellites within their line-of-sight, computation devices with built-in receivers can determine their location (longitude, latitude, and altitude/elevation) to precision within a few meters, using the Time Difference of Arrival (TDoA) method (Hofmann-Wellenhof et al. 2008). Current operational GNSS are the American Global Positioning System (GPS), the Russian Global Navigation Satellite System (GLONASS), the Chinese BeiDou Navigation Satellite System (BDS), and the European Galileo system. The Japanese Quasi-Zenith Satellite System (QZSS) currently enhances the accuracy of the GPS and plans to provide satellite navigation independent of GPS in the future (Kaplan and Hegarty 2017). The Indian Regional Navigation Satellite System (IRNSS) currently offers regional location services and plans to expand to a global version in the long term. These systems can be used to determine and track the location of vehicles during medium or long-distance transports.

3.2.2 Terrestrial Systems

There are also terrestrial systems that allow locating an object within a larger area. One example for the latter is the tracking of mobile phones through the location areas of a telecommunications network using the cell-of-origin (COO) method, where the location at a certain time is approximated by the circular area covered by the cell the device is connected to (Strang et al. 2008). In addition, the extensive installation of wireless local area networks (WLAN or Wi-Fi) and hotspots allows a more exact location based on the identification of previously mapped WLANs and hotspots.

Local positioning systems do not provide global coverage but have a limited range. Therefore, they can be used for locating a vehicle on a compound. Indoor positioning systems are optimized for use within individual buildings, where satellite signals cannot be received. They typically offer centimeter accuracy. Some provide orientation information, in addition to location information. Such systems are often used in automobile production shops, e.g., at the end-of-line. For these latter purposes, a variety of systems and technologies are in use. A well-known technology is Bluetooth Low Energy (BLE), using beacons to emit signals. The position can be computed via trilateration or via fingerprinting. Other systems are based on Long Range Wide Area Network (LoRa), Industrial, Scientific and Medical (ISM) Band, or Ultra-Wide Band (UWB) (Zekavat and Buehrer 2019).

3.3 Communication Technologies

In addition to collecting information and locating the position of vehicles, the forwarding of information is of particular importance. Accordingly, common communication technologies are presented below without any claim to completeness.

3.3.1 Wireless Local Area Network

Wireless Local Area Network (WLAN) refers to a local radio network. It is often also called Wi-Fi. In principle, one access point is sufficient for such a radio network. However, in order to cover larger areas, several access points are used. The network is operated in the free frequency band at 2.4 GHz. With commercially available transmitters, ranges of 30 to 100 meters can be achieved in the free field. With special omnidirectional antennas, 100 to 300 meters can be covered. If the WLAN is connected to the internet, data can be exchanged worldwide (Rech 2012).

3.3.2 Mobile Communications

Mobile radio is understood to mean the standards and regulations with which communication takes place. There are different technologies and systems for mobile communication. For example, the fifth-generation (5G) is currently available in Germany and enables transmission rates of up to 10 Gbit/s. The cellular network works similarly to a WLAN network (Pham et al. 2020). For example, the devices connect to an access point, and the information is transported from there to the desired recipient.

Another form of mobile communication is satellite communication. The access points are satellites in orbit. However, this type of communication is usually associated with high costs.

3.4 Technologies and Standards for Data Exchange

The Electronic Product Code Information Services (EPCIS) industry standard, which has been developed by the international non-profit organization GS1, aims to provide fine-grained data about objects, such as their position and status at a certain time, within a certain context, across companies in value-added networks

(GS1 EPCglobal 2016; Werthmann et al. 2016). Its primary function is to allow tracking and tracing of objects based on historical and current event data (Tröger 2014).

Event data describes the completion of a specific business process step acting upon one or several objects (GS1 EPCglobal 2016). It includes data about the identity of the objects related to a visibility event (what), the location of the event (where), the time and date of the event (when), and the reason or business context (why) of the event. The event data may be generated by automatic identification or localization of objects, using the already described technologies, or directly in software systems such as enterprise resource planning (ERP) systems.

Individual objects are referenced at the instance level by a globally unique Electronic Product Code (EPC) identifier, taking a Uniform Resource Identifier (URI). Examples are the Serial Shipping Container Code (SSCC) or the Global Individual Asset Identifier (GIAI). The standard provides two location types, read points and business locations, to describe the location where an event occurred. Both use a globally unique Global Location Number (GLN). To describe the event time, i.e., the date and time at which the event took place, the standard uses either local time and time zone or a globally unique UTC (Universal Time Coordinated) timestamp. The business context includes the business step in which the event occurs, the business state of the object(s) after the event, shipping and receiving parties, links to relevant business transaction documents, and instance or lot level master data.

Besides the EPCIS industry standard, another important family of independent branch standards that can be used for the global identification of logistic objects has been created by the ISO/IEC Joint Technical Committee 1 (JTC 1). JTC 1 is a joint subcommittee of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). These standards fall into three categories: A first category specifies systematic numbers for identifying objects. One example is the standard ISO/IEC 15459, which describes a globally unique identification number for identifying logistic units, the so-called Unique Transport Unit Identifier (synonymously referred to as License Plate), to identify a transport unit. The identifier consists of three components: an Issuing Agency Code (IAC), which specifies the encoding of the license plate, a Company Identification Number (CIN), which is issued by certified issuing agencies, and an individual serial number. In sum, these are encoded in a maximum of 35 ciphers or capital letters (ISO 2014). A second category includes standards for encoding data on data carriers, which can be attached to logistic units, e.g., barcodes or RFID transponders. Examples are Code 128, a set of symbols for high-density linear barcodes defined in ISO/IEC 15417:2007 (ISO 2007), and ISO/IEC 18000-63, which defines a set of parameters for air interface communications with RFID transponders at 860 MHz to 960 MHz (ISO 2015). A third category describes standards for the exchange of data between IT applications across company boundaries. These include United Nations/Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) and Extensible Markup Language (XML). UN/EDIFACT is a still widely used international standard for electronic data interchange (EDI), which provides a set of syntax rules to structure data, an interactive exchange protocol as well as standard messages for multi-country and multi-industry exchange (Kischporski 2017). XML is a mark-up language for encoding documents in a human-readable and machine-readable format.

In sum, these standards allow a comprehensive encoding of data on logistic objects, like finished vehicles, during the distribution process. Like UN/EDIFACT, some standards are dated and, while still widely used, in the process of replacement by newer standards.

4 Developed Assistance and Control Systems

In order to derive benefits from the technologies presented in sect. 3, the recorded data can be processed further and integrated into assistance and control systems. The following sub-sections give an overview of the possible functionalities and levels of automation of assistance systems. Afterwards, assistance systems and control algorithms for fully automated systems developed at BIBA are presented.

4.1 Possible Functionalities and Levels of Automation

Assistance systems can range from a simple, e.g., graphical representation of data, over recommender systems that derive recommendations for specific actions to take up to assistance systems that decide fully autonomously (Power 2002). In the context of the research project "RFID-based Automotive Network" (Werthmann et al. 2014), Werthmann et al. classified assistance systems with respect to finished vehicle logistics into five categories (Werthmann et al. 2013; Werthmann et al. 2016). The categorization ranges from low IT support and complexity of the system to high complexity and IT support (Fig. 5):

1. **Level 1: Tracking and Tracing**: Tracking and tracing consists of the two terms *tracking*—knowing where an object is and what its status is—and *tracing*—the chronological sequence of tracked events. The technology increases the transparency within the supply chain and can prevent time-consuming search activities for specific vehicles. In addition, the track and trace information can be used to estimate the scheduling of subsequent events, such as the delivery date for the customer.
2. **Level 2: Reporting**: This level comprises the transformation and aggregation of the recorded data into new key figures (key performance indicators—KPI). They can, for instance, be derived from the chronological course of events in the distribution chain and be used for the analysis and optimization of processes. Derived process times for certain work steps are an example of such key figures.

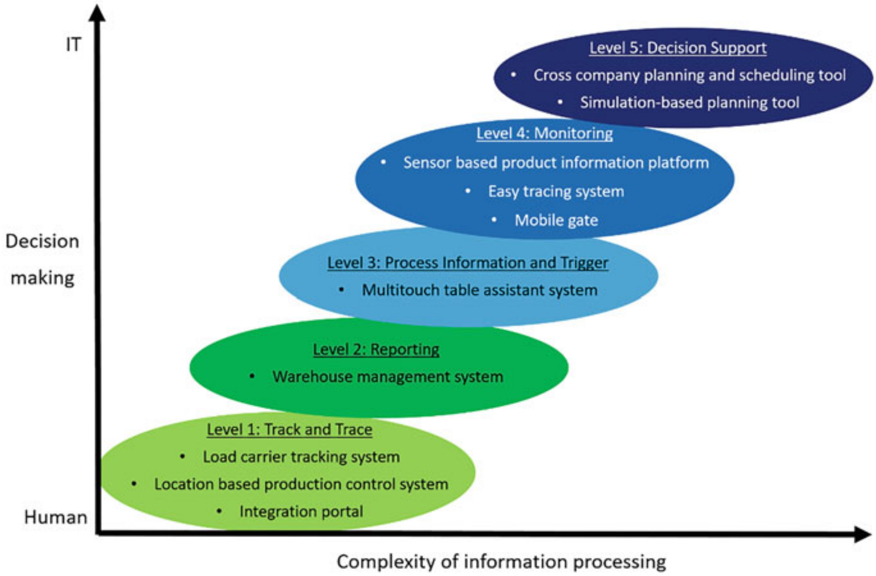


Fig. 5 Categories of functionalities of assistance systems (adapted from Werthmann et al. 2013)

3. **Level 3: Process Information and Trigger**: This level of an assistance system includes displaying process information for the user and automated triggering of subsequent processes. The triggers must be based on simple rules comparable to those that can be found in e-Kanban systems. By displaying the process information, users can adequately prepare subsequent steps, check the availability of necessary resources, and request them if necessary.
4. **Level 4: Monitoring**: This level includes monitoring processes and automated warnings when processes deviate from the target process. In the context of finished vehicle logistics, a possible function would be a warning if certain vehicles have not yet been released for loading or in case of a deviation of planned and available resources. The early detection of deviations from the plan creates scope for correcting them and prevents extensive follow-up costs.
5. **Level 5: Decision Support**: At the level of decision support, systems can weigh up concrete options for the user and present a suitable selection or recommendation. For instance, a decision support system can recommend the sequence in which technical services are applied in the rework of the manufacturer or in service centers of a vehicle compound. Furthermore, events can be triggered automatically, whereby the difference to level 3 is that the decision functions are more complex and can include artificial intelligence. For instance, several alternative options for an action can be evaluated, and the best option can be chosen. In this way, the advantages of mathematical algorithms over manual decisions can be made use of in terms of considered complexity and computation speed. Furthermore, this allows for more flexibility within the distribution chain

and prevents misdecisions. These latter systems, which no longer require any action on the user's part, are referred to in this article as **fully automated systems**.

In the following, the systems developed at BIBA and within the LogDynamics network are clustered into assistant systems and fully automated systems. Furthermore, it can be distinguished whether the systems can be applied at individual points in the distribution chain and whether they are designed to be used across several links of the distribution chain.

4.2 Assistance Systems

4.2.1 Assistance Systems at Distribution and Handling Points

Various assistance systems for automobile logistics have been developed in the past years at BIBA. Most of them are designed for use on vehicle compounds. At this point within the distribution chain, there is a particularly high risk for disruptions (Schenk and Clausen 2020), and thus, assistance systems offer a valuable potential for improvement. In the following, the developed assistance systems are presented according to their level of automation:

Load Carrier Tracking System Load carriers are used at vehicle compounds for loading high and heavy goods that cannot move on their own. The load carriers do not have fixed locations, such as a depot, on the compound. Therefore, search activities can be required to know about the status of resource availability (Scholz-Reiter et al. 2010a; Lampe 2012). The load carrier tracking system offers a user interface that displays the current locations of the load carriers on a digital map of the compound. In order to track the location, RFID, WLAN, and GPS technology are employed (Scholz-Reiter et al. 2009c; Scholz-Reiter et al. 2010b). The load carriers are equipped with RFID tags, and the tug masters that transport the load carriers are equipped with an RFID reader and a GPS receiver. When a load carrier is parked, the tug master sends its GPS position and the ID of the transported load carrier via WLAN to a database. From there, the user interface obtains the data to display the load carriers on the map. The users can apply filter functions to search for specific load carriers. In addition, the system offers possibilities to support maintenance and inventory processes and billing processes as the load carriers are owned by different players that receive compensation when their carrier is in use (Scholz-Reiter et al. 2010a; Scholz-Reiter et al. 2009a). The system by itself is a tracking system and offers potential for higher levels of assistance.

Passive Real-Time Locating System This location system is based on the Mojix system, which is a terrestrial system for both indoor and outdoor use and uses passive RFID transponders for localization. It can be deployed in technical rework processes at the manufacturer or in technical service centers on vehicle compounds.

A tracking function informs the worker in which zone (e.g., at which technical station) a vehicle is located or how busy the stations are. On the one hand, the workers save time finding a particular vehicle. On the other hand, the increase in transparency on the utilization of the technical stations allows to improve the overall capacity utilization (Werthmann et al. 2016). The overview on capacity utilization is particularly suitable for process steps that can be freely arranged regarding their sequence. A possible strategy, in this case, is always to choose that process step next for which technical stations offer the shortest waiting time. A simulation study shows that by doing so, the inventories at the stations, the throughput times, and the latter's predictability are reduced compared to the case in which the process steps keep a fixed order. The extent of this effect depends on the average utilization of the stations. The higher the utilization rate, the greater the advantage of the location system (Werthmann et al. 2012). The passive real-time location system corresponds to a level-1-assistance system. In perspective, the transparency provided by the location function can be combined with an automated control system, which detects delays at an early stage and adjusts subsequent processes accordingly (Scholz-Reiter et al. 2012b).

Warehouse Management System This software system can manage warehouses and technical services on a vehicle compound and control the movements of the vehicles on the compound. It allows tracking and tracing functions so that the vehicles' current location, status, and history are available. Furthermore, the system determines process times in order to improve future processes on the vehicle compound and thus offers reporting functions. Therefore, the system is a level-2-assistance system (Werthmann et al. 2013; Werthmann et al. 2016). The input to the system can be realized via interfaces to the "Mobile Gate" and "easy Tracing System" which are both presented below.

Mobile Gate The Mobile Gate is a gate equipped with RFID technology to identify passing vehicles (Werthmann et al. 2016). It can be set up at changing positions on a compound. This flexibility is particularly advantageous for loading or unloading ships, as they can dock in different areas. The system requires vehicles to be equipped with RFID tags. The gate then identifies them as they pass through and verifies that the vehicle is being unloaded at the correct port or loaded onto the correct ship. The gate can also generate EPCIS events stored in a higher-level information system (InfoBroker, see sub-sect. 4.2.2). In addition, the next destination of the vehicle on the compound is shown to the driver. By showing this information, the next process step, the driving process, is triggered (Werthmann et al. 2013). The system is thus a level-4-assistance system.

Easy Tracing System The easy Tracing System (eTS) is a wearable computing system used by the drivers on a vehicle compound. It consists of (i) a sensor box with an RFID reader, a Bluetooth module and a microcontroller, (ii) a mobile phone with GPS, Wi-Fi, a GPRS module, and a touch screen, (iii) a proximity sensor, (iv) an RFID antenna, and (v) batteries for energy supply (Fig. 6; Bleisteiner et al. 2014; Mrugala et al. 2008). Via the mobile phone screen, a driving order and

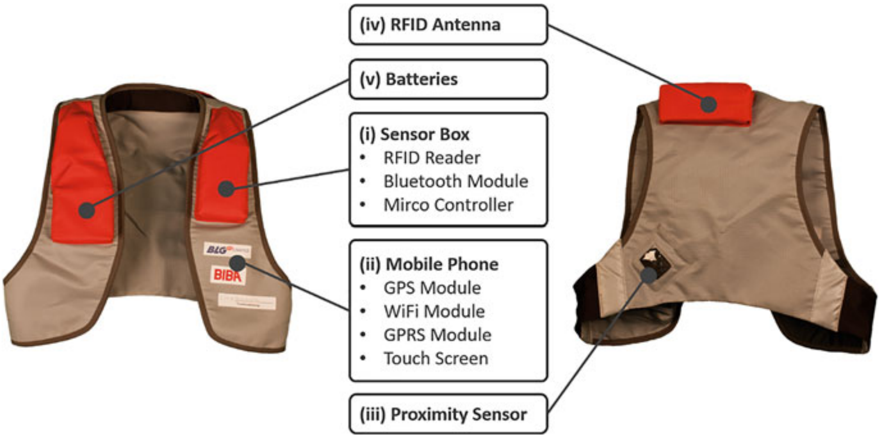


Fig. 6 Easy Tracing System

the corresponding vehicle identification number (VIN), origin, and destination are displayed to the driver. As soon as the driver sits down in the vehicle, the proximity sensor is activated, and the vehicle is identified via the RFID components. Next, it is checked whether it is the correct vehicle. When the driver leaves the vehicle, the geo-position is determined, and it is checked whether it corresponds to the correct destination. The position is then transmitted to a database via Wi-Fi (Werthmann et al. 2016). With the trigger and monitoring functions, the eTS is a level-4-assistance system (Werthmann et al. 2013).

Multi-Touch Table Assistance System This assistance system, developed by Hoff-Hoffmeyer-Zlotnik et al. (2020a), displays the current situation of a vehicle compound on a multi-touch table. It is meant to be used by planners of the compound, and the display shows a map of the compound, which indicates the level of occupancy for all parking areas. In additional information boxes, it displays the occupancy of the quays and tracks and shows status reports of current processes. Furthermore, the users can filter for vehicles assigned to certain processes, such as loading or unloading processes. These vehicles get highlighted at their position on the compound. Several planners can operate the multi-touch table at the same time. This assistance system is a level-3-assistance system due to displaying the process information.

Simulation-Based Planning Tool The planning tool is based on the same multi-touch table as described in the previous paragraph. In this setting, the vehicle compound planners can define various medium-term planning scenarios via the multi-touch table. An integrated simulation environment evaluates those in return (Hoff-Hoffmeyer-Zlotnik et al. 2017). Parameters that can be evaluated are order volume, personnel and area planning for the next few weeks to months. The tool presents the results to the planners in graphical form. The planners can retrieve the specific results for different points in time via a slide controller (Hoff-Hoffmeyer-Zlotnik et al. 2020a). As the planning tool has complex simulation software and the simulation of the scenarios goes beyond human planning capabilities, it can be seen as a Decision Support System on level 5.

4.2.2 Assistance Systems for Use Across the Distribution Chain

Cross-Company Planning and Scheduling Tool The cross-company planning and scheduling tool is a software that aims at improving coordination between compound service providers, technical service providers, and transport service providers (Scholz-Reiter et al. 2009d; Scholz-Reiter and Meinecke 2010). It involves two parts: transportation planning and order sequence planning. Based on the input of the dealer's orders, the tool proposes truck tours. For the proposition of the tours, the tool takes the transport duration, the transportation capacity, the remaining transit time, and the place of delivery into account. The transport service providers receive the suggestions, and they can revise them if necessary. The final formation is forwarded back to the planning and scheduling tool. Subsequently, it proposes a processing sequence for the technical service centers. The sequence is planned via backward scheduling, and the aim is to ensure that the vehicles of the same tour leave the technical service center at the same time. This also allows to lower the size of the dispatching area, the trucks' waiting time, and the duration of the loading processes. The technical service centers receive the proposals and can edit them in case needed. This completes the planning process. As such, the cross-company planning and scheduling tool is a level-5-assistance system. A simulation study shows an improvement of all considered key factors compared to the situation where the tool is not used. The tool lowers the average utilization of disposition area and increases the average load factor, the average on time delivery, the average number of carrying truck tours, and the average number of transported vehicles (Scholz-Reiter and Meinecke 2011).

InfoBroker The InfoBroker is a system that allows for the cross-company exchange of event data in a standardized form. The event data are, for instance, RFID-generated events along the distribution chain. The underlying standard is EPCIS. The event data gives information on which object was identified at which time at what place for what reason (Werthmann et al. 2016). The system allows to integrate local InfoBroker instances into networks. At the same time, an important feature is a security system that prevents unauthorized access to sensitive data. The companies are categorized according to their role, e.g., manufacturer, logistic service provider, or dealer. Depending on their role, they can receive certain event data and are in charge of supplying certain event data (Brandwein et al. 2013). By regulating the exchange of information based on the company's role, it is ensured that companies only receive information that they are authorized for.

Moreover, the data exchange is limited to information relevant for decision-making. Within the companies' IT structure, the information of the InfoBroker can be forwarded to assistance systems that further process the data according to the companies' needs (Werthmann et al. 2013). Therefore, the InfoBroker is not an assistance system by itself but can act as a valuable basis for such systems.

Integration Portal The integration portal is a track and trace system that spans across the entire distribution chain and is based on the InfoBroker described in the previous paragraph (Werthmann et al. 2013). Filter functions allow for requesting status information to vehicles, transports, processes, and installed parts. The corresponding vehicles are displayed on a map.

Sensor-Based Product Information Platform This platform allows monitoring the status and logistical quality of products and product components within supply and distribution chains. The platform integrates geolocation, mobile sensors, and telecommunication technologies to record the geolocation of transports as well as quality-relevant environmental influences on the products, e.g., humidity and temperature. Cloud-based digital services analyze the collected sensor data and issue warnings in case of undesirable events, like quality impairments or delivery delays (Teucke et al. 2018; Servos et al. 2020). To facilitate exchange of quality related event-based data across company boundaries, expressions for the exchange of sensor and quality data have been added to the EPCIS standard. The platform is a level-4-assistance system.

4.3 Control Algorithms for Fully Automated Systems

4.3.1 Distinguishing Features

Several control algorithms have been developed and evaluated in the past years at BIBA that are the basis for control systems on vehicle compounds. These control algorithms use complex decision rules to make fully autonomous decisions about ongoing processes. The objectives of the algorithms are the allocation of parking spaces, the processing order within technical service centers, and the allocation of driving and transport orders to drivers and shuttles. The approaches follow decentralized, centralized, or hybrid control paradigms. Decentralized approaches have the advantage of being able to make use of swarm intelligence. That means they can achieve good system behavior via simple decision rules evaluated within each logistic object (Hongler et al. 2010). Centralized methods can either follow simple rules, such as FIFO. In those cases, they are believed to be outperformed by decentralized approaches. On the other hand, centralized approaches allow for the implementation of complex optimization problems and thus for optimal solutions (Hongler et al. 2010). Hybrid approaches strive to combine the advantages of both.

4.3.2 Control Algorithms at Distribution and Handling Points

Agent-Based Assignment System for Parking Spaces The approach for agent-based assignment of parking spaces on a vehicle compound is a decentrally organized negotiation mechanism. The negotiation happens between the logistic objects, i.e., the vehicles and the parking spaces. The goal of the vehicles is to find the parking space with the shortest possible total travel time to the parking space and from there to the subsequent destination. The goal of the parking spaces is the highest possible occupation. The system can be implemented using RFID, GPS, and IC technologies where the components are to be combined similarly to the easy tracing system described in sub-sect. 4.2 (Böse 2012). A simulation study compares the agent-based assignment mechanism to a central control approach in which the vehicles are assigned to a parking space according to fixed priority rules whereby the priority is not necessarily linked to the length of the travel distances. Using the agent-based negotiation approach, the travel times per vehicle decreased by 26 seconds. In sum, this results in a time saving of 112 working days per year for the compound (Böse and Windt 2007; Böse et al. 2008).

Hybrid Control System for Parking Space Allocation The hybrid control system is an extension of the agent-based approach described above. The hybrid control combines two modules: on the one hand, a module for the agent-based, decentralized parking space allocation, and on the other hand, a module for centralized control of parking space allocation that acts according to priority rules (Scholz-Reiter et al. 2011). The advantage of the decentralized control module is short travel times. The advantage of the centralized control module is a higher degree of sorting within the parking areas enabling shorter delivery times for subsequent processes. A switching unit that activates the two control modules depending on the situation on the compound aims to combine the advantages of both approaches. A simulation study shows that the hybrid control actually achieves similarly low values for the vehicles' travel times as the decentralized module. In terms of sorting, the hybrid control system achieves values between centralized and decentralized control modules but closer to that of centralized control (Scholz-Reiter et al. 2012a).

Autonomous Control Methods for Process Control Simulation studies investigate the throughput of vehicles through a vehicle compound and compare it for different autonomous control methods (Windt et al. 2010; Becker and Windt 2011). The studies view two processes on the compound: the selection of a parking space (storage process) and the subsequent optional passage through a technical service center (manufacturing process). Four variants of the scenario are set up, and control algorithms from the domain of autonomous control, self-control, multi-agent-based methods, smart systems, and decentralized control are used. These are compared against two benchmark methods. The standard benchmark method provides a fixed order of priorities regarding the selection of parking spaces, whereby the priority differs for vehicles with and without options for technical services. The technical service stations are visited in a fixed order. As a second benchmark scenario, a random method is implemented. Parking spaces and the order in which the vehicles pass through the technical service center are randomly selected according to a uniform distribution. The key performance indicators are average time to parking and average parking utilization (storage process) as well as average total treatment time of a vehicle and average due date reliability (manufacturing process). The simulation studies show that the autonomous methods can outperform the benchmark methods but that none of the autonomous methods dominates all other methods overall key figures and scenarios.

Combined Optimal Control System for Order Assignment and Shuttle Routing The combined optimal control system for order assignment and shuttle routing is designed for the utilization on a vehicle compound where shuttles are operating to transport handling employees to their subsequent order. A control algorithm assigns driving orders to handle employees and transport orders to shuttles. Furthermore, it coordinates the routing of the shuttle buses. The control algorithm is based on a central control approach that acts via combined optimization. The aim is to achieve the fastest possible processing of all driving orders. The technical implementation is based on a server and smartphones with GPS positioning, an indoor positioning system and WLAN. The server hosts the control algorithm and communicates with the smartphones via WLAN. An app informs the handling employees and shuttle drivers about their next orders. Furthermore, a GPS module on smartphones locates the vehicles when they are parked at their destination. A geofencing system is employed to check whether the correct destination has been reached. If not, the handling employee is informed about the mismatch (Hoff-Hoffmeyer-Zlotnik et al. 2020a). Inside parking racks, where GPS positioning does not work due to the metallic surroundings, a Wi-Fi Round Trip Time-based indoor positioning system is set up (Jathe et al. 2019). Overall, two routing options are available for routing the shuttles: branch-and-bound routing and FIFO routing. The variants differ in terms of waiting times for the shuttles and travel time in the shuttles. However, both offer similar efficiency for processing transfer orders (Hoff-Hoffmeyer-Zlotnik et al. 2020b). Current research works on extending the control algorithm by artificial intelligence modules. The aim is that a combination of critic and actor neural networks make suitable initial guesses for the solution of the optimization problem such that computation time can be saved.

Compound Control System The compound control system is based on the combined optimal control system for order assignment and shuttle routing described above and is currently under development at BIBA. The aim is to extend the system to make it usable for driving orders that start or end within deep-sea ships or enclosed train wagons. Within these modes of transportation, usual Wi-Fi coverage and GPS signal are not available due to the metallic surrounding. The research aims to develop a system that allows for data transfer into and out of the ships and trains such that status reports and new orders can be transmitted. The aimed-for technology is ad-hoc mesh networks based on LoRa Lite technology that spans between the drivers and shuttle drivers within the ship/train. At least one driver or shuttle driver that is in contact with the Wi-Fi network on the compound must always be included to connect the mesh network to the main system. Moreover, two mobile applications are under development in this project: Firstly, an application that integrates the stevedores into the compound control systems and allows them to share the stowage plan with the system and request specific vehicles in real-time during the process of loading. Secondly, a mobile application that integrates special cases into the compound control system.

The following section provides BIBA's vision of a transparent finished vehicle distribution chain under adaptive planning and control. The section describes how the presented assistance and control systems contribute to this vision and gives a retrospective on BIBA's research contribution in this field.

5 Towards Automobile Logistics 4.0

The vision developed during BIBA's research activities in vehicle distribution processes emphasizes two main points: (1) full transparency of the vehicle distribution chain and (2) fully adaptive planning and control of the vehicle distribution chain. After describing both, the section will present a retrospective on the general trajectory of BIBA's research and future prospects.

5.1 Vision 1: Fully Transparent Vehicle Distribution Chain

Digital, Industry 4.0-related technologies can increase transparency for the various processes along the vehicle distribution chain. BIBA's vision of a fully transparent finished vehicle distribution chain that integrates the described technologies, systems, and solutions and makes use of them in a comprehensive manner is shown in Fig. 7: Auto-identification systems based on technologies like RFID or barcode can identify each vehicle at each station in a timely manner. Individually adapted location systems using either satellite-based or terrestrial technologies can determine the vehicles' current locations, either during a long-distance transport or within a limited area, like a vehicle compound or a production shop. These systems can transfer their data automatically to databases that allow data storage of the vehicles during each process step. Standardized information transfer mechanisms, like EPCIS and the InfoBroker, enable efficient data transfer across company boundaries. As a result, more, better, and timelier information is available for every operation.

This transparency gain results in processes that are more efficient and less wasteful. For example, a local information system, like the Easy Tracing System, can immediately give a warning signal when a driver tries to park a vehicle at a wrong (unauthorized) parking area on a vehicle compound. Moreover, the system stores the chosen parking position. Consequently, the frequent searches for vehicles on compounds that require high workforce capacities and thus cause high costs can be substantially reduced.

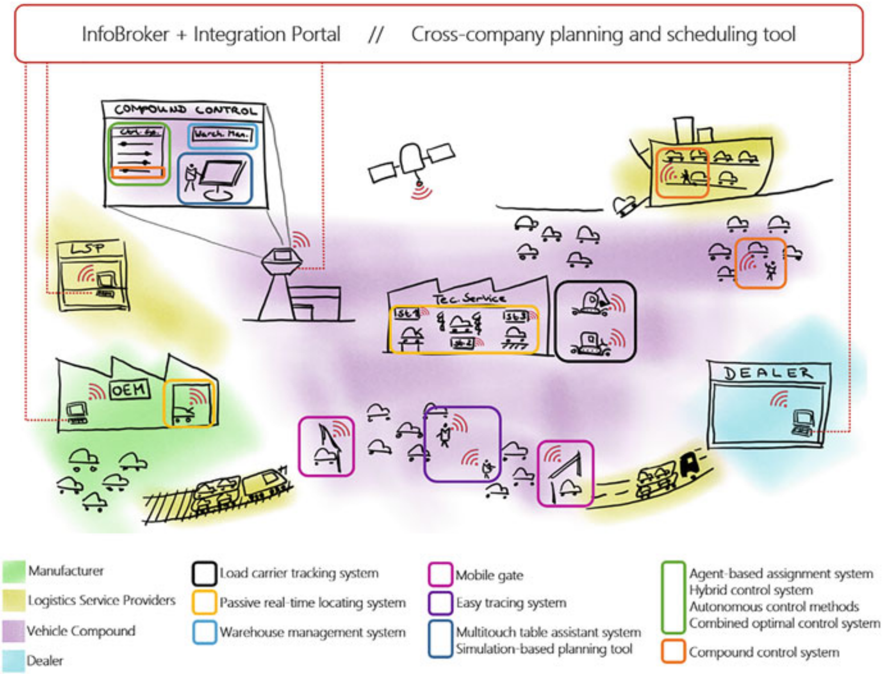


Fig. 7 Vision of a transparent finished vehicle distribution chain

In addition, the increased transparency allows better planning and control of logistics operations. If crucial data is missing, it cannot be considered for planning and control. If data is wrong or obsolete, planning or control will be erroneous as a result. In any case, planning and control will be sub-optimal.

5.2 Vision 2: Adaptive Planning and Control of Vehicle Distribution Chains

In addition to benefits from increased transparency, planning, and control processes themselves become better by using digital technologies. Automated planning and control processes supported by modern information processing technologies and based on suitable methods and algorithms can handle higher system complexity than previous methods primarily based on manual execution. Planning horizons can be extended further into the future, while planning cycles can be run more often. As a result, planning horizon and planning cycle times can be adapted to specific problems and situations.

Concerning the comprehensiveness of planning and control systems, two different developmental trajectories become possible: On the one hand, seamless information flow between different processes enables the collaborative and parallel planning of the different links and the integrated planning across the whole distribution chain. This global planning of the distribution chain, as exemplified by the cross-company planning and scheduling tool, allows a better overall planning solution than the isolated planning of each link with only rudimentary consideration of adjacent processes.

On the other hand, digital technologies allow delegation of previously centralized planning or control functions to individual elements of a system. Here, the concept of autonomous control allows, for instance, the decentralized assignment of parking spaces, or service times, based on negotiations of agents representing logistic objects, like vehicles, parking areas, and service stations. Such autonomous decision-making can relieve overall planning systems of unnecessary detail planning and make vehicle distribution systems more robust to better cope with high complexity and dynamics. Furthermore, hybrid approaches, such as the Hybrid control system for parking space allocation, allow combining the strengths of both autonomous control and central planning and control.

5.3 Retrospective on the Research and Future Prospects

The general course of research started with addressing more straightforward problems and developing basic but adequate solutions and proceeded towards more complex problems and solutions. For instance, the design of assistance systems started with the simpler Load carrier tracking system, using RFID, WLAN, and GPS technology to display locations of load carriers on the compound. The easy Tracing System, as an intermediate system, included a sensor box with an RFID reader, a Bluetooth module and a microcontroller, a mobile phone (with GPS, Wi-Fi, a GPRS module, and touch screen), a proximity sensor, an RFID antenna and batteries for energy supply. The most recent technical solution for locating vehicles, shuttles, and drivers on a vehicle compound combines a server, GPS positioning, an indoor positioning system, WLAN, and smartphones.

A similar course can be observed with regard to planning and control solutions. These started from a relatively simple, agent-based parking space allocation to cars, then added additional autonomous control solutions for service center operations. The most recent planning and control solutions integrate online simulation to evaluate medium-term planning alternatives or combine order assignment and shuttle routing using optimization methods. In order to achieve the best possible overall solution, these recent developments are not based on autonomous objects processing local information but rather on using ICT to provide accurate information on the current state of the terminal as a whole. In addition, increased emphasis is put on making the automatically computed data available to human planners via sophisticated interfaces for human–machine interaction.

Due to their particular importance as focal points within the distribution network and their sheer size, vehicle compounds have so far been the object of most attention and research activities. Consequently, most of the described systems and solutions address vehicle compound operations. This is particularly the case for assistance systems that allow higher automated or more efficient local processes. In addition, this may set a topic for future research activities to transfer assistance systems developed for vehicle compounds to use at other links of the distribution chain, such as the smaller storage yards of dealers. Several solutions are available for improving the planning and control across the distribution chain. Thus, future research can address the challenge to integrate local assistance systems with these more comprehensive solutions.

The newly developed systems have to be integrated into the existing, very heterogeneous IT infrastructures of the OEMs, LSP, vehicle compound operators, and dealers. The large variety of different and often mutually incompatible IT systems renders integration across the vehicle distribution chain very challenging. A first solution to this problem requires proper, standardized interfaces for data exchange. Often, a standardized interface for data exchange, such as the InfoBroker, may be sufficient. In other cases, closer functional integration may be required. Another possibility may be the usage of cloud computing systems.

The final motive of introducing digital systems is economic profit. Several potentials to leverage economic profit by reducing costs or increasing customer satisfaction have already been mentioned. In addition, increased digitalization offers potentials for new business models that may open up new revenue streams for the companies involved in vehicle distribution. For instance, the increased data available may allow fact-based evaluation and benchmarking of different transport routes and relations, allowing the long-term strategic optimization of vehicle distribution networks. Another business model is using transparency data to improve insurance and security systems. The tracking and tracing systems allow a fact-based and automated allocation of responsibility for vehicle damage cases. This may reduce the high costs associated with damage classification and allocation. In addition, vehicle distributors may sell data to insurance companies, which allows the definition of insurance conditions better adapted to actual risks.

6 Summary and Outlook

BIBA and its project partners have developed a broad range of planning and control methods and assistance systems in the past years that span across the entire finished vehicle distribution chain and lend valuable support to the stakeholders of the distribution process. With minor adaptations, the majority of the methods and systems can become useful for different stakeholders and stages within the distribution network. Many methods and systems developed for use on vehicle compounds can handle operations of OEMs or dealers as well. Often, simplified versions should be sufficient for these stakeholders.

With continuing advances in mobile computation and location technologies, further progression of the developed systems is possible (Fig. 8). Current research

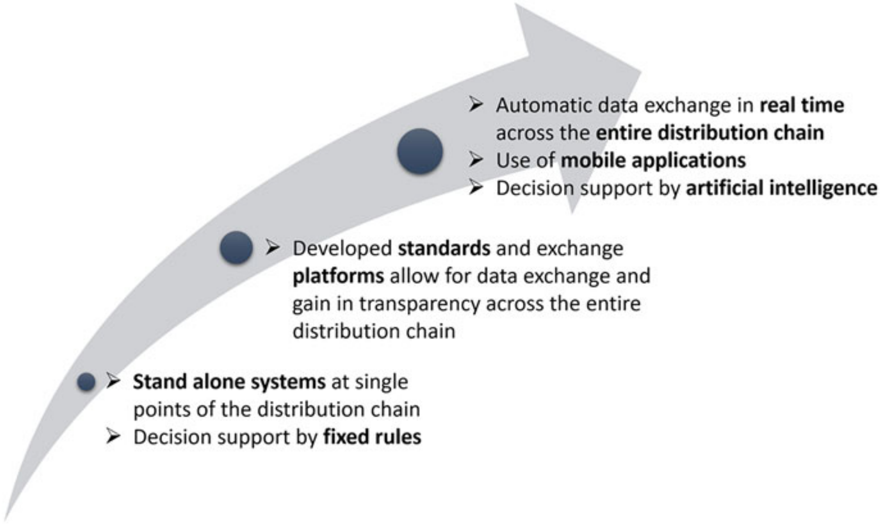


Fig. 8 Line of development in transparency and decision support systems within the finished vehicle distribution chain

shows that existing systems can be extended to additional groups of workers or further processes, such as stevedores or loading and unloading vehicles. Moreover, mobile applications offer significant potential for the digitalization of processes. This prevents, on the one hand, media disruptions and associated errors. On the other hand, data can be shared in real-time with adjacent processes and, if handy, with the entire distribution chain.

The advances in the field of artificial intelligence offer further potential for even better support. Assistance systems and fully automated systems can be improved for better recommendations based on more intelligent planning and control methods and decision rules that no longer need to be implemented manually. This opens up a new era of assistant and recommender systems, which can adapt to dynamic changes in vehicle distribution services.

Acknowledgments The research presented in this chapter was funded by German Research Foundation (DFG) as part of the Collaborative Research Center 637 “Autonomous Cooperating Logistic Processes—A Paradigm Shift and its Limitations,” by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the projects “ProKon—Use of innovative ICT technologies for process control in cargo and load carrier management at seaports,” “LogPro—Logistical planning and control systems in RoRo and ConRo ports,” “RAN—RFID-based Automotive Network,” and “SaSch—Digital Services for Shaping Agile Supply Chains,” and by the German Federal Ministry of Transport and Digital Infrastructure (BMVI) within the projects “ISABELLA—Automobile logistics in sea and inland ports: interactive and simulation-based operation planning, dynamic and context-based control of device and load movements” and “ISABELLA2—Automobile logistics in sea and inland ports: Integrated and user-oriented control of device and load movements through artificial intelligence and a virtual training application.”

References

Ambe, I.M., Badenhorst-Weiss, J.A.: Strategic supply chain framework for the automotive industry. African J Bus Manag. **4**, 2110–2120 (2010)

Becker, T., Windt, K.: A comparative view on existing autonomous control approaches- observations from a simulation study. In: Hülsmann, M., Scholz-Reiter, B., Windt, K. (eds.) Autonomous Cooperation and Control in Logistics, pp. 275–289. Springer-Verlag, Berlin (2011)

Bleisteiner, T., Ramirez, A., Werthmann, D.: Die RAN-Erfassungsklassen. In: Transparenz in globalen Lieferketten der Automobilindustrie, pp. 185–203. Publicis Publishing, Erlangen (2014)

Böse, F.: Selbststeuerung in der Fahrzeuglogistik. Modellierung und Analyse selbststeuernder logistischer Prozesse in der Auftragsabwicklung von Automobilterminals. Gito Verlag, Berlin (2012)

Böse, F., Windt, K.: Autonomously controlled storage allocation on an automobile terminal. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics – the Impact on Management, Information and Communication and Material Flow, pp. 351–363. Springer, Berlin (2007)

Böse, F., Piotrowski, J., Scholz-Reiter, B.: Autonomously controlled storage management in vehicle logistics – applications of RFID and mobile computing systems. Int J RF Technol Res Appl. **1**, 1–20 (2008)

Brandwein, D., Werthmann, D., Scholz-Reiter, B.: Infobroker und RFID-Technik in der multimodalen Fahrzeugdistribution. In: 18. Magdeburger Logistiktage. pp. 53–61. (2013)

Dias, J.C.Q., Calado, J.M.F., Mendonça, M.C.: The role of European «ro-ro» port terminals in the automotive supply chain management. J. Transp. Geogr. **18**, 116–124 (2010)

Filz, B.: Kennzahlensystem für die Distribution: Modell für kleine und mittlere Unternehmen. TÜV Rheinland, Köln Verl (1989)

Finkenzeller, K.: RFID Handbuch. Grundlagen und praktische Anwendungen von Transpondern, kontaktlosen Chipkarten und NFC. Hanser, München (2015)

Fischer, T.: Multi-Agenten-Systeme im Fahrzeugumschlag – Agentenbasierte Planungsunterstützung für Seehafen-Automobilterminals. Deutscher Universitäts-Verlag, Wiesbaden (2004)

Genc, E., Duffie, N., Reinhart, G.: Event-based supply chain early warning system for an adaptive production control. Procedia CIRP. **19**, 39–44 (2014)

Görges, M., Freitag, M.: Modeling autonomously controlled automobile terminal processes. In: Proceedings of the Hamburg International Conference of Logistics (HICL 2019). pp. 185–214. (2019a)

Görges, M., Freitag, M.: Dynamisierung von Planungsaufgaben auf Automobilterminals. Ind 40 Manag. **35**, 23–26 (2019b)

GS1 EPCglobal: EPC Information Services (EPCIS) Standard. Release 1.2, Ratified, Sep. 2016. GS1 AISBL, Brussels, Belgium. https://www.gs1.org/sites/default/files/docs/epc/EPCIS-Standard-1.2-r-2016-09-29.pdf. (2016). Accessed 25 May 2021

Gudehus, T.: Logistik. Grundlagen, Strategien, Anwendungen. Springer, Heidelberg (2010)

Gudehus, T.: Logistik 1. Grundlagen, Verfahren und Strategien. Springer Vieweg, Berlin (2012)

Hammer, M., Champy, J.: Business reengineering : die Radikalkur für das Unternehmen. Campus-Verl, Frankfurt/Main (2003)

Haubrich, H.: Optimierung der Fahrzeugdistribution mit Software der INFORM GmbH. In: Göpfert, I., Braun, D., Schulz, M. (eds.) Automobillogistik, pp. 287–304. Springer Gabler, Wiesbaden (2017)

Hermes, A.M.: Modellbasierte Bewertung von Potenzialen einer distributionsorientierten Programm- und Reihenfolgeplanung in der Automobilindustrie. Praxiswissen Service, Dortmund (2011)

Herold, L.: Kundenorientierte Prozesssteuerung in der Automobilindustrie. Deutscher Universitäts-Verlag, Wiesbaden (2005)

Hoff-Hoffmeyer-Zlotnik, M., Schukraft, S., Werthmann, D., Oelker, S., Freitag, M.: Interactive planning and control for finished vehicle logistics. In: Jahn, C., Kersten, W., Ringle, C.M. (eds.) Digitalization in Maritime and Sustainable Logistics, pp. 77–93. epubli, Berlin (2017)

Hoff-Hoffmeyer-Zlotnik, M., Schukraft, S., Sprodowski, T., Freitag, M., Dunekacke, D., Zeitler, V., Görges, M.: Gestaltung von Mensch-Technik-Schnittstellen für die Digitalisierung von Planungs- und Steuerungsprozessen auf Automobilterminals. In: Freitag, M., Barthelmey, A. (eds.) Mensch-Technik-Interaktion in der digitalisierten Arbeitswelt, pp. 263–282. GITO mbH Verlag, Berlin (2020a)

Hoff-Hoffmeyer-Zlotnik, M., Sprodowski, T., Freitag, M.: Revisiting order assignment problems in a real-case vehicle compound scenario. In: Colloquium in Information Science and Technology, CIST. pp. 395–400. (2020b)

Hofmann-Wellenhof, B., Lichtenegger, H., Wasle, E.: GNNS – Global Navigation Satellite Systems. Springer, Wien, New York (2008)

Holweg, M., Miemczyk, J.: Logistics in the "three-day car" age – assessing the responsiveness of vehicle distribution logistics in the UK. Int. J. Phys. Distrib. Logist. Manag. **32**, 829–850 (2002)

Hongler, M.O., Gallay, O., Hlsmann, M., Cordes, P., Colmorn, R.: Centralized versus decentralized controlA solvable stylized model in transportation. Phys A Stat Mech its Appl. **389**, 4162–4171 (2010)

International Organization for Standardization. ISO/IEC 15417:2007. Information technology — Automatic identification and data capture techniques — Code 128 bar code symbology specification. Online: https://www.iso.org/standard/43896.html. (2007). Accessed 10 June 2021

International Organization for Standardization. SO/IEC 15459-1: 2014. Information technology — Automatic identification and data capture techniques — Unique identification — Part 1: Individual transport units. Online: https://www.iso.org/standard/54779.html. (2014). Accessed 10 June 2021

International Organization for Standardization. ISO/IEC 18000-63: 2015: Information technology – Radio frequency identification for item management – Part 63: Parameters for air interface communications at 860 MHz to 960 MHz Type C. Online: https://www.iso.org/standard/63675.html. (2015). Accessed 10 June 2021

Jathe, N., Lütjen, M., Freitag, M.: Indoor positioning in car parks by using Wi-fi round-trip-time to support finished vehicle logistics on port terminals. IFAC-PapersOnLine. **52**, 857–862 (2019)

Jesse, R., Rosenbaum, O.: Barcode: Theorie, Lexikon, Software, 1. Aufl. Verl Technik. Berlin. (2000)

Kagermann, H., Wahlster, W., Helbig, J.: Umsetzungsempfehlungen fuer das Zukunftsprojekt Industrie 4.0. acatech. Deutsche Akademie der Technikwissenschaften e.V, München (2013)

Kaplan, E.D., Hegarty, C.J.: Understanding GPS/GNSS. Principles and Applications. Arttech House, Boston, MA (2017)

Kischporski, M.: EDI – Digitalisierung und IT-Wertbeitrag konkret umgesetzt. Springer Gabler, Wiesbaden (2017)

Klug, F.: Logistikmanagement in der Automobilindustrie – Grundlagen der Logistik im Automobilbau. Springer Vieweg, Berlin (2018)

Lampe, W: nachhaltig bewegen. Nachhaltigkeitsbericht 2012. Bremen. (2012)

Lödding, H.: Handbook of Manufacturing Control: Fundamentals, Description, Configuration. Springer, Berlin (2011)

LogDynamics. Bremen Research Cluster for Dynamics in Logistics. http://www.logdynamics.de/logdynamics.html?&L=1. (2021)

Monostori, J.: Supply chains robustness: challenges and opportunities. Procedia CIRP. **67**, 110–115 (2018)

Mori, S., Suen, C.Y., Yamamoto, K.: Historical review of OCR Research and Development. Proc. IEEE. **80**, 1029–1058 (1992)

Mrugala, D., Ruthenbeck, C., Scholz-Reiter, B., Lang, W.: Smart jacket as a computing system for automobile warehouse logistics. In: European Conference & Exhibition on Integration Issues of Miniaturized Systems. Brüssel, pp. 492–495. (2008)

Musa, A., Gunasekaran, A., Yusuf, Y.: Supply chain product visibility: methods, systems and impacts. Expert Syst. Appl. **41**, 176–194 (2014)

Nait-Sidi-Moh, A., Bakhouya, M., Gaber, J., Wack, M.: Geopositioning and Mobility. ISTE, London (2013)

Negri, E., Fumagalli, L., Macchi, M.: A review of the roles of digital twin in CPS-based production systems. Procedia Manuf. **11**, 939–948 (2017)

Pfohl, H.-C.: Logistiksysteme. Betriebswirtschaftliche Grundlagen, 8th edn. Springer, Berlin (2010)

Pham, Q.V., Fang, F., Ha, V.N., Piran, M.J., Le, M., Le, L.B., Ding, Z.: A survey of multi-access edge computing in 5G and beyond: fundamentals, technology integration, and state-of-the-art. IEEE Access. **8**, 116974–117017 (2020)

Power, D.J.: Decision Support Systems. Concepts and Resources for Managers. Quorum Books, Westport (2002)

Rech, J.: Wireless LANs. 802.11-WLAN-Technologie und praktische Umsetzung im Detail. Heise, Hannover (2012)

Ross, D.F.: Distribution Planning and Control: Managing in the Era of Supply Chain Managemen. Springer, Boston, MA (2015)

Ruthenbeck, C., Lappe, D., Lampe, W.: Informationsmanagement in der Automobillogistik. Ind Manag. **26**, 9–13 (2010)

Schenk, L.: ActiveBarcode. Barcode Software seit 1994. https://www.activebarcode.de/ (2021)

Schenk, A., Clausen, U.: Creating transparency in the finished vehicles transportation process through the implementation of a real-time decision support system. IEEE Int Conf Ind Eng Eng Manag. **12**, 1017–1021 (2020)

Scholz-Reiter, B., Meinecke, C.: Planungsmethode für die unternehmensübergreifende Auftragsabwicklung. Ind Manag. **26**, 39–42 (2010)

Scholz-Reiter, B., Meinecke, C.: Nutzen von Informationstransparenz im Seehafenterminal. Prod. Manag. **16**, 56–58 (2011)

Scholz-Reiter, B., Böse, F., Dizdar, C., Windt, K.: Transportorientierte Reihenfolgeplanung – Verbesserte Abstimmung zwischen Produktion und Distribution. PPS Manag. **13**, 15–17 (2008)

Scholz-Reiter, B., Böse, F., Teucke, M., Virnich, A.: Ladungsträgermanagement auf RoRo-Terminals. Ind Manag. **25**, 25–28 (2009a)

Scholz-Reiter, B., Frazzon, E.M., Makuschewitz, T.: Integration von Produktions- und Transportsystemen entlang globaler supply chains. Ind Manag. **25**, 31–34 (2009b)

Scholz-Reiter B., Isenberg M. A., Virnich, A., Özsahin, M. E.: Innovative load carrier management solution for seaport terminals – based on positioning, identification and communication technologies. In: Proceedings of the 5th international Congress on Logistics and SCM Systems (ICLS 2009). pp. 638–644. (2009c)

Scholz-Reiter, B., Meinecke, C., Ruthenbeck, C.: Development of a cross-company planning and scheduling method for the supply chain management of RoRo and ConRo-Ports. In: 1st International Conference on Logistics and Transport. pp. 1–15. (2009d)

Scholz-Reiter, B., Isenberg, M. A., Virnich, A., Özsahin, M. E.: Automatic load carrier management on RoRo terminals. In: 2010 Annual Conference of the International Association of Maritime Economists. pp. 2–19. (2010a)

Scholz-Reiter, B., Özsahin, M. E., Schweizer, A., Isenberg, M. A.: Combined information and communication technologies for location and status transprancy on Ro-Ro terminal. In: 12th Int. Conference on Harbor, Maritime and Multimodal Logistics Modeling and Simulation, HMS 2010. pp. 127–132. (2010b)

Scholz-Reiter, B., Lappe, D., Toonen, C., Ruthenbeck, C.: Development of a hybrid control approach for automotive logistics based on close to real-time information management. Int J Syst Appl Eng Dev. **5**, 545–552 (2011)

Scholz-Reiter, B., Lappe, D., Ruthenbeck, C., Toonen, C.: Hybride Steuerung logistischer Prozesse. wt Werkstattstech. **102**, 240–245 (2012a)

Scholz-Reiter, B., Werthmann, D., Lappe, D., Otterstedt, N.: Ortungsgestützte Steuerung von Qualitätssicherungsprozessen. Ind Manag. **28**, 22–26 (2012b)

Schuh, G.: Distributionslogistik. In: Schuh, G., Stich, V. (eds.) Logistikmanagement. Handbuch Produktion und Management 6. Springer, Berlin (2013)

Schukraft, S., Oelker, S., Werthmann, D., Freitag, M., Görges, M., Gencer, E., Malek, A.: Interaktive Planung und Steuerung für den Automobilumschlag. Ind 40 Manag. **33**, 11–14 (2018)

Seeck, S.: Erfolgsfaktor Logistik. Klassische Fehler erkennen und vermeiden. Gabler, Wiesbaden (2010)

Servos, N., Liu, X., Teucke, M., Freitag, M.: Travel time prediction in a multimodal freight transport relation using machine learning algorithms. Logistics. **4**, 1–22 (2020)

Strang, T., Schubert, F., Oberweis, R.: Lokalisierungsverfahren. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Oberpfaffenhofen. (2008)

Teucke, M., Broda, E., Börold, A., Freitag, M.: Using sensor-based quality data in automotive supply chains. Mach. Des. **6**, 1–22 (2018)

Tröger, R.: Supply Chain Event Management – Bedarf, Systemarchitektur und Nutzen aus Perspektive fokaler Unternehmen der. Universität Leipzig, Modeindustrie (2014)

Werthmann, D., Lappe, D., Otterstedt, N., Scholz-Reiter, B.: Ortungsgestütztes Produktionsleitsystem für die Nacharbeit bei Automobilherstellern. wt Werkstattstech. **102**, 114–119 (2012)

Werthmann, D., Brandwein, D., Ruthenbeck, C., Scholz-Reiter, B.: Standardized information exchange in automotive distribution processes. In: 2013 IFAC Conference on Manufacturing Modelling, Management and Control. pp. 700–705. (2013)

Werthmann, D., Bender, M., Engelhardt, P., Winkler, M., Parlings, M., Reinhardt, S.: Von der Idee zur Umsetzung – ein Leitfaden. In: Lepratti, R., Lamparter, S., Schröder, R. (eds.) Transparenz in globalen Lieferketten der Automobilindustrie, pp. 60–116. Publicis Publishing, Erlangen (2014)

Werthmann, D., Brandwein, D., Ruthenbeck, C., Scholz-Reiter, B., Freitag, M.: Towards a standardised information exchange within finished vehicle logistics based on RFID and EPCIS. Int. J. Prod. Res. **55**, 4136–4152 (2016)

Windt, K, Becker, T, Kolev, I: A comparison of the logistics performance of autonomous control methods in production logistics. In: Proceedings of the 43rd CIRP International Conference on Manufacturing Systems. Vienna, pp. 576–583. (2010)

Zekavat, S.A., Buehrer, R.: Handbook of Position Location. IEEE Press, Piscataway (2019)

15 Years of Intelligent Container Research

Reiner Jedermann and Walter Lang

Abstract Food losses in the cool chain, which are mostly caused by temperature deviations, can be reduced by remote monitoring of transport conditions. The project ‘Intelligent Container’ was begun 15 years ago to provide the necessary sensor system, communication and automated evaluation of data. If transport and delivery planning are adjusted according to the actual quality or the predicted remaining shelf life, more products arrive with sufficient quality at the customers. This paper summarizes the project results and highlights current trends in industrial application and research, such as commercial remote container monitoring and standards for data exchange, sub-GHz communication, the increasing availability of biological and computational fluid dynamics models and digital twins. Open research topics include the development of specialized sensors. To overcome obstacles hindering the industrial application of sensor quality monitoring, we suggest a gradual approach, with lower company resources required for the first action points. Food losses can be reduced, even if the complete system, including permanent remote access and adaptive stock rotation, is not applied.

1 Introduction

Fifteen years ago, we set our research focus on the ‘Intelligent Container’ as a means to mitigate food losses caused by temperature abuse and deviating environmental conditions during logistic operations. The Intelligent Container provides remote monitoring to detect transport problems as early as possible and predict resulting effects on food quality, using the length of the remaining shelf life as a scale. The container operates as an autonomous system by predicting the shelf life condition

---

R. Jedermann · W. Lang (✉)
Institute for Microsensors, –Actuators and –Systems (IMSAS), University of Bremen, Bremen
Germany
e-mail: rjedermann@imsas.uni-bremen.de; wlang@imsas.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_11

in which the product will arrive at the destination point. If it is foreseeable that the product will not arrive with sufficient quality, a warning message will be generated.

The loss of food quality is a gradual process, often hidden to visual inspection until it is too late (Nunes et al. 2014), i.e. manual quality inspections at the end of a transport process will often not reveal the effects of previous harmful temperature conditions. Later in the distribution centre, at the retailer or even at the private customer, the product will decay long before its printed use-by date.

The means to reveal such invisible losses is to calculate the shelf life based on the product's temperature and sensory history. This approach can be well integrated into warehouse management. According to the first-expires-first-out (**FEFO**) principle, products with a low shelf life are assigned to nearby destinations, and products with high shelf lives are held back for longer transport routes. By using this intelligent stock rotation, the effects of deviating transport conditions, often not visible from the outside, can be mostly compensated, and more products arrive at the customer with a quality level above the acceptance threshold. A survey in (Jedermann et al. 2014a) showed that this approach could reduce losses of highly perishable products by 8% to 14% of the total transported volume.

1.1 Outline

In this paper, we will provide a short overview of the project history and present the most important findings. International research and industrial applications have further advanced since the completion of our field tests on ocean vessels in 2013. We summarize these current developments and analyze obstacles that still hinder the broad application of the FEFO concept. Finally, we list three subsequent action points by which companies can begin to reduce losses and mitigate temperature problems, even if only a small budget and resources are available.

1.2 Project History

There are several reasons to fix the birth of the Intelligent Container to the year 2006, although the term was used previously and the Collaborative Research Centre SFB637 on autonomous logistic processes began two years earlier. A first study with three industrial partners on the magnitude of temperature deviations began in 2006, founded by an internal budget of the Bremen Research Cluster for Dynamics in Logistics (LogDynamics). In the same year, a small-scale demonstrator was presented at the Hannover trade fair (Fig. 1). The webpage http://www.intelligentcontainer.com/ was registered, and the first journal paper was published (Jedermann et al. 2006).

The transfer of research results to the industry was started in 2008 by a sub-project of the SFB637 with three companies: Cargobull, a provider of truck

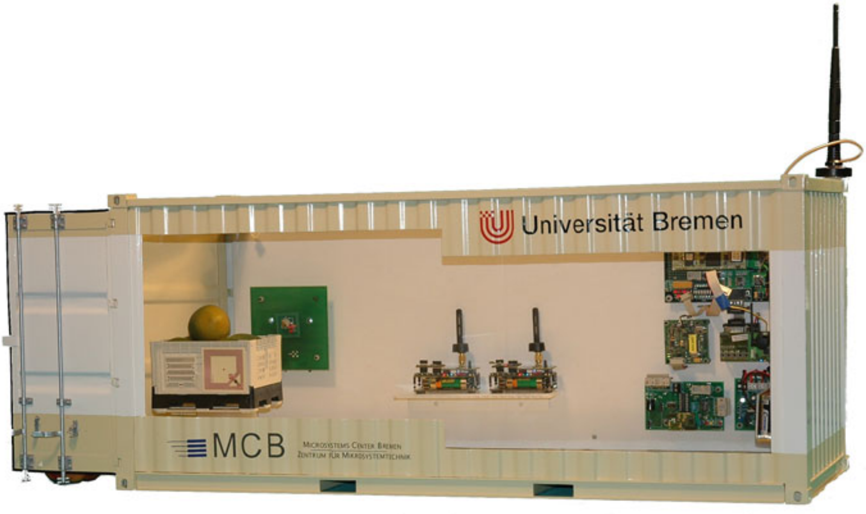


Fig. 1 Model container (1:8) for demonstration purposes (Jedermann et al. 2007a). The initial concept included the following steps: New freight items are scanned by RFID (left). Sensor nodes supervise the environmental conditions (middle). The processing module (right) calculates shelf life prediction

Fig. 2 Full-size prototype of the intelligent container and research team of Bremen University Bremen

telematics, Rungis Express, a delicacy food trading company, and Dole Fresh Fruit as an importer for bananas. Tests for the wireless sensor monitoring of bananas in ocean containers commenced in 2009 (Fig. 2). In 2010, the cooperation broadened to an Innovation Alliance with 15 industrial and 6 research partners, leading to the publication of a dedicated theme issue on ‘Intelligent Food Logistics’ in 2014 (Jedermann et al. 2014a). Subsequent projects covered the analyses of airflow

Fig. 3 Main elements of the cooling chain for bananas (see text for details)

conditions in cold storage warehouses by wireless anemometers (Hartgenbusch et al. 2020), the detection of ethylene by micro-fabricated gas sensors (Sklorz et al. 2012a,b) and the detection of fungi spores (Papireddy Vinayaka et al. 2016).

1.3 The Banana Chain

Most of our experimental work was dedicated to bananas. We carried out five test transports with a sensor and communication-equipped container. Bananas were harvested green in Central America, packed into boxes, stacked to pallets, stowed to reefer containers, trucked to the harbour and shipped to Europe (Fig. 3). Bananas were stowed ‘warm’ to the container. Cooling began when the container was connected to a power supply at the harbour in Central America. In Europe, the pallets were stored in a climatized warehouse for a few days or weeks and trucked to a ripening facility in Germany. There, the conversion from starch to sugar with the outside colour change from green to yellow was initiated by exposure to ethylene gas. Commercial ripening is only feasible with completely green bananas; otherwise, it is not possible to achieve uniform good quality. Bananas must arrive at the ripening facility before their green life ends. We monitored the process from stowing the pallets to the container until the completion of the ripening process.

2 Findings

Our project results cover various fields, from sensor development, electronics and communication to biological research and logistic models.

2.1 Omnipresence of Temperature Deviations

Although we have only anecdotical evidence, we tend to say that, 'Wherever you measure temperature, you find deviations'. During our very first temperature data logger test in a refrigerated delivery truck, we observed a strange spike. Later, we learned that the cooling unit had run out of oil, but the driver was able to fix it after an hour.

Particularly in smaller companies, awareness of maintaining correct temperatures without interruption is lacking. A supplier of meat agreed on participating in a temperature monitoring study. Thus, he proved himself guilty of not cooling during transport from slaughterhouse to depot. He was banned from the suppliers' list by the import company with the consequence that we had to adapt microbiological models to another meat product in our project.

The average temperature over two weeks' transportation of bananas in ocean containers was between 2.5 °C and 4.9 °C above the setpoint during our tests (Jedermann et al. 2019). The deviations of the average container temperature are overlaid with temperature heterogeneity in vertical and horizontal directions (Fig. 4). The vertical temperature mostly follows a regular pattern with an increase of approximately 1 °C from the lowest to the sixth box layer on the pallet. Further, upward to the eighth top layer, temperature slightly decreases due to additional cooling by the horizontal flow of return air.

Deformations of containers and pallets of several centimetres from an ideal cuboid shape cause an irregular distribution of gap diameters and airflow. Thus, the horizontal temperature pattern is rather erratic. For new, up-to-date equipment, we observed a horizontal variation 1.5 °C, with higher values for older containers.

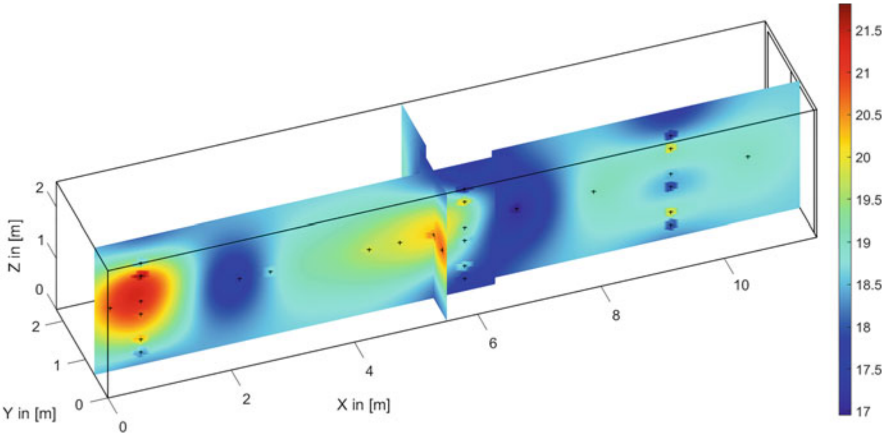


Fig. 4 Average temperature during an ocean transport of bananas in March 2011 with a temperature setpoint of 14.4 °C. Crosses mark the sensor positions. Temperature on the slice planes was interpolated by Kriging

A 12-year-old container required more than double the time to cool to a stable temperature, and after two weeks' cooling, it was still 1.5 °C above the setpoint (Jedermann et al. 2014c).

Temperature fluctuations over the length of the container can be analyzed using a variogram (Jedermann et al. 2011). Measurement points with a distance larger than the variogram range only contribute to the common average but have no further statistical relation. To capture the spatial temperature profile, the sensor distances must be less than the variogram range. For banana containers, we estimated a variogram range between 1.13 and 1.65 meters (Jedermann et al. 2011).

Other authors found temperature variations of similar magnitude, e.g., Pelletier et al. (2011) observed a temperature increase of 4.5 °C after 5 days of truck transportation. According to a survey by Ndraha et al. (2018), temperature limits were overstepped in 13.6% to 58% of transports.

During a truck transport of grapes in Brazil, a 4 to 11 °C temperature variation was observed (Oliveira et al. 2021). Humidity varied from 60% to 98%. The shelf life differed by a factor of almost two depending on the calculated mass loss.

Human and machinery failure can cause much higher temperature abuse than the above-described variations on regular transports. The most important reasons for claims are, according to (Castelein et al. 2020), reefer unit malfunction and power failures, followed by failing to cool products before stowing (hot stuffing), incorrect temperature setpoint, poor packing and poor cargo quality at loading. Delays at transhipment, leaving pallets without cooling at the platform, is another serious cause of temperature variation (Nunes et al. 2014), particularly due to an insufficient workforce at peak seasons.

2.2 Necessity of Sub-GHz Communication and Gateway

Translating a laboratory system into a real-world test is always cumbersome. One of our most challenging experiences was that 2.4 GHz wireless sensors do not work if placed between bananas. A later theoretical model, tuned by our experimental data on signal strength, showed that 2.4 GHz radio waves are attenuated by −62 dB per meter by packed bananas (Jedermann et al. 2014b). Communication was only possible over a maximum distance of 50 cm. This problem can be solved by switching to sub-GHz radio modules. For 866 MHz, the attenuation drops to −8.3 dB/m and for 433 MHz to −2.2 dB/m. Sub-GHz sensors with Lora, Dash7 and proprietary protocols were successfully tested to penetrate a stack of apple crates over 2 to 5 meters (Jedermann et al. 2018).

Not only water-containing food products but also the metal hull of containers hinder radio communication. Penetration of a metal box requires high radio power, which is not feasible for battery-powered devices. The recommended solution consists of a gateway that connects the low-power internal wireless network with external communication. This gateway can also be connected to the power supply of the cooling unit and has, therefore, a higher energy budget for communication.

The freight supervision unit in our Intelligent Container links the internal wireless sensor network with external Wi-Fi, GSM and Iridium satellite communication. It also provides a platform to calculate shelf life models and generate warning messages. The functionality of the system could be verified in several field tests.

2.3 *Shelf- and Green-Life Models*

Understanding the complete kinetics of the chemical and enzymatic reactions causing quality and shelf life loss is not practical. Laboratory data are typically limited to external attributes such as colour, firmness and weight loss. The dependency of these attributes to the sample's temperature history can be modelled by using a simplified approach:

The product has an initial budget of shelf life after harvest. A certain amount is subtracted every day, depending on the temperature. At optimal conditions, one day of shelf life is lost per day of transportation. The loss per day (LPD) is accelerated by deviating transport conditions. The simplest form of the LPD is given by an exponential function of temperature, or it can be more accurately calculated by the Arrhenius Law for reaction kinetics (Jedermann et al. 2014c). By switching from a day-wise calculation to an integral form, arbitrary temperature curves can be used as input for the shelf life model as in our model for the green life of bananas (Jedermann et al. 2014c).

The model assumes that a temperature deviation at the beginning of the cool chain has the same effect as a deviation of the same height and length at the end of the chain or two separate deviating events with both half lengths. In the latter case, more temperatures fluctuations can lead to condensation and, thus, increased bacterial growth.

More elaborate models attempt to predict the change of quality attributes by using temperature-dependent differential equations for the concentration and activity of enzymes, e.g., models developed by the Frisbee project for apples and packed cooked ham (Gwanpua et al. 2015). The model for apples also includes the influence of air humidity on moisture loss. The identification of 10 or even more model parameters sets high demands on the extent of laboratory experiments.

As a compromise, bacterial growth can be described by common pattern functions, such as the Gompertz function applied in our project for meat products (Bruckner et al. 2013).

2.4 *Ethylene Detection*

Ripening fruits can communicate. For example placing a ripe banana between green bananas initiates the ripening of the other bananas. The trigger is ethylene, a gaseous ripening hormone. The effect works in both directions: climacteric fruits react to ethylene, and ethylene is produced by the fruit after the ripening has begun. High-resolution measurement of ethylene will be useful for the management of fruit logistics because it can directly detect the ripening status of the fruits in the container.

In gas chromatography, the adsorption and desorption of a gas in a chromatographic column are used to separate gases. To achieve resolution in the ppm range, the ethylene gas is accumulated in a pre-concentrator before chromatographic measurement (Janssen et al. 2014). In the last 20 years, small and autonomously working micro-gas chromatographic systems for ethylene have been developed. The fluidic system is miniaturized down to a size of approximately 10 cm. The critical parts, i.e. the chromatographic column and the pre-concentrator, are made using silicon micro-technology (Sklorz et al. 2012a). Ethylene can be detected in the 10-ppb range (Zaidi et al. 2017). Since 2015, it has been possible to replace expensive silicon with robust and low-cost materials, such as 3D polymer printing for the columns (Lucklum et al. 2015) and ceramics for pre-concentrators (Zaidi et al. 2018).

An alternative method to sense ethylene is to measure the specific infrared light absorption of ethylene molecules at 10.6-μm wavelength by NDIR (non-dispersive infrared spectroscopy) (Sklorz et al. 2012b). The method is highly robust and easier to implement than gas chromatography, but the resolution is only in the 10-ppm (part per million) range. Extending the resolution to the ppb range using advanced systems and pre-concentrators is possible but has yet to be developed for container applications (Popa and Udrea 2019; da Silveira Petruci et al. 2014).

2.5 Models for Heat Removal and Generation

The cooling unit of a container must remove heat from three different sources: compensation for isolation losses of the container walls, remaining field heat from the product in case of insufficient or missing pre-cooling and heat generated by the product itself. The later heat load can become the highest factor for climacteric fruits, such as bananas. Biological processes and respiration continue after harvest. The conversion from starch to sugar produces large amounts of the heat up to 200 W/ton during ripening, whereas green unripe bananas produce only 20 W/ton to 50 W/ton depending on temperature and atmosphere conditions (Jedermann et al. 2014c). If the produced heat extends the local heat removal by cooling of typically 30 W/ton to 100 w/ton, a hotspot is generated with uncontrollable temperature rises.

The local heat removal per box is subject to variations in packing and diameters of air gaps between the pallets. Our model for bananas (Jedermann and Lang 2014) enables to describe heat removal and generation by two separate proportional factors. If the bananas begin to ripen, the latter factor can increase after a few days or weeks. The first factor can be identified by the temperature data from the first week of container transport. Thereafter, the actual heat generation can be estimated, and the formation of a hotspot can be detected in an early state.

This monitoring system enables the artificial ripening of bananas to be carried out directly in the transport container. The cargo hold is exposed to ethylene, and the temperature is slightly increased to initiate ripening. Thereafter, the temperature is gradually decreased to maintain the ripening process at a moderate speed and heat generation without overstepping the cooling capacity.

2.6 Case Study on Cool Chain Logistics

The possible reduction of food losses by the Intelligent Container was evaluated in a case study on bananas (Haass et al. 2015). Temperature-related losses were simulated for a logistic network between a harbour in Central America and ten customers in Europe. After arrival in one of three European ports, containers were to be reassigned to another customer depending on their predicted remaining green life. Simulation parameters were tuned according to an analysis of the existing logistic network and order quantities provided by one of our project partners. The assumption of probabilistic distribution of banana quality from an earlier study (Lütjen et al. 2013) was replaced by a green life model providing a more accurate prediction of losses.

The scenario with conventional containers, without information about green life changes, resulted in 4.27% losses of the total transported volume. Reassignment of deliveries according to the actual green life prediction reduced the losses by 0.37%. Accepting short delays in delivery enables more flexible transport reassignments. If additionally, causes of temperature deviations are detected earlier by the Intelligent Container, losses can be reduced by 0.88% at a maximum.

2.7 Detection of Fungus Spores

Mould is substantial danger in the transport of food (Pitt and Hocking 2009). To detect fungus, spores are caught on a nutrient medium surface, where they can grow. The growing mould changes the medium's pH, and this can be detected by impedance measurement (Papireddy Vinayaka et al. 2016). Alternatively, the pH change can be used to trigger a colour change detected by a camera.

Classically, the growth of fungal colonies on a culture medium is observed using a light microscope by a trained human observer. This observation has successfully been performed by computer vision and image recognition (Tahir et al. 2018). Fungal spores are immobilized, grown, an optic micrograph is taken, and the fungus colonies are identified and counted by a convolutional neural network (Blank et al. 2016).

2.8 Difficulties in Quality Measurement and Prediction for Green Bananas

Bananas have the highest seaborne reefer cargo volume per year for any fruit (Castelein et al. 2020). The fact that all partners in the cool chain were in the hand of one company simplified our tests from the farm to the ripening facility in Germany. Nevertheless, the achievable savings by FEFO are generally quite limited for bananas. The simulation showed that 0.88% of total volume losses could be avoided by adaptive order assignment. This value is substantially less than the previously mentioned savings for other highly perishable food products (8–14%).

Green bananas are a highly robust product if handled with care and stored at the correct temperature. Under a controlled atmosphere, they can last for 2 months or longer in a warehouse. The artificial ripening process evens out all differences in green banana ages. Only turners with visible colour changes must be excluded from ripening.

However, there are other points contributing to the 'nastiness' of bananas: a green banana is just a green banana. There is no outside indicator to determine the age of a green banana—neither colour, firmness, nor respiration activity. A laboratory examination of enzyme concentrations is far beyond feasible in transport supervision. The only way to predict green life is based on temperature history, supported by additional data on the atmosphere, harvest age and growing conditions.

Green life models can provide recommendations regarding which pallets should be prioritized for further processing, but they are far from predicting the precise day on which the banana will turn from green to yellow. The green life model has a typical standard deviation of ±5 days (Jedermann et al. 2014c). This biological variance adds to the effect of temperature variation inside one container, being the larger one of the two (Fig. 5).

Although FEFO brings only limited advantages for bananas, our project results demonstrated that the banana chain profits largely from remote quality monitoring. Maladjusted temperature set points, power and machinery failures are detected immediately. In case of general quality problems, the farm can be informed before the next ship is sent on its way, transporting containers with most likely the same quality problem. Replacement deliveries can be organized in time; this ability is particularly crucial if boxes with a quality problem were branded for a specific customer. The temperature setpoint of the cooling unit can be fine-tuned according to the actual box temperature. The effect of alternate packing and stowage schemes was verified by our tests. Finally, we could demonstrate that container ripening is feasible based on detailed sensor control and modelling.

Fig. 5 Monitoring temperature changes as cause for green life variations. Wireless temperature/humidity sensor (middle box) and temperature data logger (corner box). Transparent boxes were only used for demonstration purposes

3 Current Developments and Trends

'Food cold chain management' is a growing research area according to a review by Shashi et al. (Shashi et al. 2020) of 1189 journal papers. The number of published papers doubled within 4 years to 216 papers in 2019. The authors identified four major research clusters. The first cluster a) is centred around RFID technologies and shelf life management. The other three clusters cover non-technical topics, such as b) production, harvesting and transport planning, c) causes and management of postharvest waste and d) methods and quality standards for the identification of problematic areas.

The limitation of cluster a) to RFID is that it appears not to represent the full bandwidth of current research and industrial applications. In the following section, we introduce communication technologies, standards, models and processing platforms that go beyond our initial concept for the Intelligent Container.

3.1 Communication

Communication is the most advancing aspect of remote cool chain control. Maersk began to equip their reefer containers with remote container monitoring (**RCM**) based on GSM communication in 2016 (Zarkani and Rasmussen 2016). More than 400 ocean vessels were equipped with a satellite dome to forward GSM data packets from the containers. In 2019, the system was extended by the 'Captain Peter' App (Avery 2019), giving all their customers instant and easy access to data of container location, temperature setpoint and cooling engine status. Nevertheless, the system is focused on monitoring those parameters that are the responsibility of the transport operator, i.e. the correct setup and operation of the cooling unit and timely delivery, but not its content status. An interface to wireless sensors inside the cargo was announced but hardly applied in practice.

LoRaWAN (long-range wide-area network) allows communication with low data rates in license-free sub-GHz frequency bands. We demonstrated that LoRa radio signals could penetrate 4.5 meters of fruits with only moderate packet losses of 1.5% (Jedermann et al. 2018), making it a highly suitable candidate for communication with packed food products inside a container. However, LoRaWAN has been, so far, mostly applied to the outside communication of containers, e.g., on the vessel or in harbour settings, where cellular networks are not available (Avery 2020).

Besides the communication range, the network capacity is a crucial factor, such as in a harbour setting with thousands of containers. According to (Gambiroža et al. 2019), the capacity of the LoRaWAN network is still an open research issue, depending on the assignment of data rates, number of end devices and external disturbances in the frequency range. They summarize the results of 10 capacity studies. A network with 8000 devices can, for example transmit 500 packets/min, which is far more than sufficient to transmit the hourly status data of containers.

Narrowband Internet-of-Things (NB-IOT) enables cellular communication with an extremely low energy budget. Loggers have a battery life of 8 years (Ako Electromecànica 2020). But they face the same problem as other GSM loggers [e.g., (MOST Tech Sweden AB 2021)] in that the metal hull of the container and signal attenuation by water-containing products hinder communication with the next cell tower at some kilometres' distance.

A more detailed survey of available RCM and wireless data logger hardware can be found in our earlier publication (Jedermann et al. 2017).

3.2 Standards

In the Intelligent Container, we adapted standards such as XML data structures, IPv6 and the Constrained Application Protocol (CoAP) (Jedermann et al. 2014b), mixed with proprietary protocols, particularly in the lower protocol layers. In recent years, wireless sensor networks have moved towards widely accepted standards, such as the LoRaWAN protocol. RFID Sensors can use available data fields in the EPC Gen-2 protocol to transmit temperature data.

The need for trans-company interoperability has been recognized by ocean carriers. The Digital Container Shipping Association (DCSA) was founded in 2019 with the goal to define open and freely accessible standards for the global container shipping industry. Up to now, their member list includes 9 of the top 10 ocean carriers. Their track and trace (T&T) standards were published on SwaggerHub and GitHub (DCSA 2021).

3.3 Modelling

The biological modelling of the quality behaviour of food products is still an active research area. The FRISBEE project published six detailed shelf life models (Bruckner et al. 2013). The University of Tasmania, Australia, has offered a website to access over 100 computational biological models to predict the growth of various food pathogens by the partly commercial CB-Premium service (https://www.cbpremium.org/) since 2018 (Tamplin 2018).

The growing number of freely available models simplifies the adaptation of FEFO concepts. However, even if a model for a particular product is already available, the model must be fine-tuned to local varieties, weather and growing conditions in order to provide an accurate prediction of shelf life variations as a function of the sensory production and transport histories. Biological experiments are still necessary, although in a reduced number.

Another area of growing research interest is the computational fluid dynamics (CFD) modelling of packing, for example to optimize their layout and the location and size of vent holes (Ambaw et al. 2017). The complete simulation of a loaded container is still beyond current computational capacities. Modelling is only feasible with some simplifications, such as replacing product pallets and air gaps with porous media (Moureh et al. 2009) or modelling only empty containers (Getahun et al. 2018). Such CFD models are useful for optimizing packing and stowage layouts and detecting general airflow and ventilation problems causing insufficient cooling. However, the actual temperature distribution largely depends on factors that are hardly measurable or predictable, such as varying gap diameters (Jedermann and Lang 2017), as well as container and pallet deformations.

3.4 Modelling Platforms, Cloud Computing and Digital Twins

When we began in 2008 with the initial tests on remote truck monitoring, our partner company was proud to have negotiated a GSM flat rate for 400 kByte per month. The limited communication volume drove the need for intelligent data processing directly inside the means of transportation. The automated detection of critical events was one of the key concepts of the Intelligent Container with the idea to reduce external communication, mostly to warning messages on such events.

Today, the inclusive volume of data tariffs has increased by a factor of 1000 or even more. Edge processing—as implemented in the Intelligent Container—is still useful for robustness in case of communication failures but no longer absolutely necessary. The trend has moved toward cloud computing and data analysis on the server level.

Our first concepts were based on mobile and autonomous software agents (Jedermann et al. 2007b), which made decisions on behalf of the freight owner, e.g., assessment of quality risks and, in case of detected problems, searching for an alternate destination with a shorter transport route.

Decision support tools (DSC) (Mack et al. 2013) are another term to describe the data processing by an Intelligent Container. The DSC analyzes temperature and additional sensor data by shelf life models and generates warnings and recommendations regarding how to adjust transport planning.

Today, the focus has broadened from isolated shelf life and temperature distribution models to more holistic approaches. **Digital twins** pursue to provide a detailed and accurate representation of real-world objects, such as a pallet of food products or other transport entities. Digital twins combine physical properties, such as size and weight, information from an electronic waybill and sensory data with extensive modelling tools. Digital twin platforms provide standard interfaces to collect data from multiple sources and interlink different types of models, such as CFD, shelf life, weather and traffic models.

Digital twins have been applied in structural health monitoring and smart manufacturing (Lu et al. 2020) and recently also in the supply chain (Defraeye et al. 2021). A special challenge is the modelling of uncertainty. Whereas mechanical systems have mostly well-known and accurate models, biological models have a high level of unknown input factors, side-effects and biological variance. These uncertainties must be represented, processed and evaluated in the digital twin platform.

Machine learning has become a hot topic in recent years, driving the idea of replacing cumbersome laboratory experiments on shelf life models with autonomous learning based on data from remote transport monitoring. Enthusiastic concepts often underestimate the size of the necessary training database and related costs. The training data must contain sufficient instances of all combinations of input factors, e.g., temperature, humidity, mechanical damages and scars during packing. Otherwise, the machine learning overfits to single deviating events. The dogma of garbage in garbage out is also true for machine learning (Camacho et al. 2018). Furthermore, supervised learning requires target output values, i.e. an evaluation of the actual product quality as a result of its transport history. The generation of such data requires manual laboratory analysis of the samples. A simple yes/no value if there had been a subsequent customer complaint provides only coarse information.

4 Conclusions and Action Points

The availability of technical components for the Intelligent Container has largely increased in recent years. New reefer trucks and containers are mostly sold with built-in remote monitoring facilities. The numbers of available wireless sensors systems, data loggers and published shelf life models are also growing.

Pilot studies have shown that losses of perishable products can be reduced by 8% to 14% of the total transport volume by scheduling deliveries according to the FEFO principle (Jedermann et al. 2014a). Nevertheless, FEFO has yet to find its way into practice to any significant degree.

4.1 Obstacles for FEFO Implementation

The obstacles that hinder FEFO's implementation must be questioned. One point is that FEFO does not fit every product and every chain. Less sensitive products with long shelf lives profit less from FEFO. Rescheduling of deliveries is mostly feasible for high-volume products with several alternate customers. If there are no products with a higher shelf life available in the warehouse, the retailer would be more likely to risk providing low-quality products than having empty shelves.

More severely, missing links on both technical and operational sides obstruct the commercial application of the FEFO concept:

Missing Link between Internal and External Communication The large-scale application of RCM by Maersk and the growing number of companies offering wireless data loggers show that there is a business case both for external and internal monitoring of containers with perishable products. However, interfaces that link these two system layers are still missing.

The manufacturers of wireless data loggers provide their own cloud platforms instead of integrating into existing RCM systems. Offering 'data as a service' is a more lucrative business model than merely selling the hardware.

Missing Cost Sharing Models for Sensor Monitoring Cool chains are mainly divided into different operators (Jedermann et al. 2017) with different interests, quality management targets for their services, profits from increased shelf life and responsibilities and costs for contributing to remote quality monitoring. The one who pays for sensors and communication tends not to be the one who ultimately profits, e.g., the producers must place sensors in the product packing but have no direct benefit. Logistic service providers for reefer trucks and containers must install the communication system to prove that their equipment is running at the correct temperature without interruption. However, this can also backfire in case of problems and raise customer complaints. The distribution centres profit from FEFO by reducing customer complaints and in-warehouse losses, mostly without physical contact to the sensor hardware. The retailers must collect the sensors. They receive products with a more predictable average shelf life and profit from higher customer satisfaction.

The fragmented chain makes it impossible to create a business case for the individual operators, who must be linked not only on the logistical side but also in the goal to minimize losses at the end of the chain and share related costs.

4.2 Recommended Practical Actions

Food losses by temperature abuse are often significant and cannot be accepted in a world of limited environmental resources, although the obstacles and investments are often too high to implement FEFO in a specific cool chain. Therefore, we suggest a set of three subsequent actions (Jedermann et al. 2019). Even if only the first action can be implemented due to limited budget and resources, losses can be reduced substantially.

Action 1: Offline Data Logger Study The first and most important action identifies and mitigates cool chain ruptures. Offline, low-cost, reusable temperature loggers are packed to the products in repeated tests to analyze the chain's performance. No dedicated communication infrastructure has to be installed at this stage. Such a study often reveals avoidable gaps in the cool chain, e.g., power interruptions during harbour handling, extended periods on loading platforms without cooling, wrong temperature setpoints and high-temperature variation inside the cargo hold due to insufficient air circulation by careless stowing.

Weak spots in the chain can be identified and corrected, e.g., excluding older reefer containers, which had more than twofold temperature deviations compared with up-to-date equipment in our tests (Gwanpua et al. 2015). The effect of correct pre-cooling at the farm or production site on the temperature performance of the subsequent transport is also often under-estimated. The rule-of-thumb that '1 hour delay in pre-cooling costs 1 day of shelf life' is mostly true, according to (Pelletier et al. 2011).

Action 2: Detection of Actual Problems by Live Data Action 2 switches from occasional offline data collection to regular real-time or live access, at least at some control points, such as arrival at the harbour or distribution centre. Permanent live monitoring requires more investments for mobile communication but also provides increased benefits. Power failures and defective cooling equipment, which are one major cause of cool chain raptures, can be detected immediately. In general, the earlier information arrives, the more time there is to plan and correct subsequent steps, e.g., prioritizing container handling in the harbour according to their degree of temperature deviation, ordering a replacement delivery in case of a large excess of temperature limits and informing the producer about possible problems.

Action 3: Intelligent Stock Rotation by FEFO If the product's temperature history is available, even as the logistical process continues, losses can be further reduced by adaptive delivery planning and stock rotation according to the FEFO concept. Products with low remaining shelf life are assigned to the shortest transport routes. Necessary investments include the analysis of the product-specific relation between temperature and shelf life. Existing models can serve as a starting point after some fine-tuning. Based on collected data from regular transport monitoring and detailed examination product quality for some samples, the model can be later refined.

4.3 Research on New Sensor Types

‘Temperature’ is the environmental condition with the highest impact on fruit quality. The accuracy of shelf life prediction and the early detection of transport problems will be improved if other sensor types are added, e.g., atmosphere conditions such as humidity, O_2 and CO_2 concentration. Sensors for other important contributing factors to transport quality and safety are still an open research topic:

Development of New NDIR Sensors Experience has shown that micro-gas chromatographic systems are difficult to implement for autonomous operations. Hence, it seems advisable to improve non-dispersive infrared measurement (NDIR) for ethylene detection. Today, NDIR resolves the lower ppm range, whereas for containers, a ppb resolution would be necessary. Recently, new NDIR systems using hollow waveguide technology have been demonstrated (da Silveira Petruci et al. 2014). They require much less gas, which would allow combining an NDIR sensor with a micromachined pre-concentrator. This way, theoretically, a resolution of 50 ppb is possible. Furthermore, using two different infrared wavelengths, a combined system, which could also measure ethylene and fumigation gases such as methyl bromide or phosphine, will be possible.

Identification of Pest Insects in the Container Cargo pest insects are a major risk in overseas transport. Countries invest significant efforts into control and law enforcement to prevent invading species. Once an insect species has invaded, it is almost impossible to suppress it. A sad example is the progress of the Chinese rose beetle, which invaded Hawaii. An extensive biological control program had no success in stopping its spread (Stanaway et al. 2001). It is recommended to develop a system that can not only detect the presence of insects in a container but also identify them and provide a warning if dangerous species are found. One possibility might be using adhesive traps and pattern recognition.

Summary Various research efforts have been carried out to reduce food losses in the last two decades. The detection of temperature abuse and the FEFO concept are two of the main contributions. The technical basis has already been provided by the Intelligent Container and supported by recent developments in commercial RCM and sensor hardware. Further research is still necessary to foster action three by additional pilot studies, which are the development of dedicated shelf life models, more detailed quantification of possible loss reductions by the FEFO concept, and research in new sensor types.

Acknowledgements The initial research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 'Autonomous Cooperating Logistic Processes', followed by the industrial cooperation project 'The Intelligent Container' supported by the Federal Ministry of Education and Research, Germany, under reference number 01IA10001. The project COOL for monitoring cold storage warehouses and the project Camsense for particle measurement in the air were supported by the Federal Ministry for Economic Affairs and Energy based on a decision by the German Bundestag.

The Joint Project MaUS—An Autonomous Microreactor System for Detection of Mould Contaminations—was founded by the German Federal Ministry of Education and Research. The development of an in-situ ethylene measurement system was founded by AiF-IGF.

We wish to thank our research partners, particularly Ulrike Praeger and Martin Geyer from the Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Potsdam and Judith Kreyenschmidt from the Institute for Animal Sciences (ITW), Bonn. Among the several industrial partners, we particularly thank Dole for their cooperation for more than 7 years in various projects. We also thank Mike Nicometo, a former member of the Cool Chain Association, for his advice on the project and for co-editing a theme issue on intelligent food logistics.

References

Ako Electromecànica. Case Study: Multinational Transport (Refrigerated Div.). https://www.ako.com/products/wireless-temperature-moisture-monitoring/nb-iot-temperature-humidity-logger/ (2020)

Ambaw, A., Mukama, M., Opara, U.L.: Analysis of the effects of package design on the rate and uniformity of cooling of stacked pomegranates: numerical and experimental studies. Comput. Electron. Agric. **136**, 13–24 (2017). https://doi.org/10.1016/j.compag.2017.02.015

Avery, P.: Maersk takes next step with reefer monitoring. In: WorldCargo News, England, https://www.worldcargonews.com/news/news/maersk-takes-next-step-with-reefer-monitoring-61619 (2019)

Avery, P.: ORBCOMM rolls out LoRa on vessels. In: WorldCargo News, England, https://www.worldcargonews.com/news/orbcomm-rolls-out-lora-on-vessels-6532 (2020)

Blank, R., Vinayaka, P.P., Tahir, M.W., Yong, J., Vellekoop, M.J., Lang, W.: Comparison of several optical methods for an automated fungal spore sensor system concept. IEEE Sensors J. **16**(14), 5596–5602 (2016). https://doi.org/10.1109/JSEN.2016.2567538

Bruckner, S., Albrecht, A., Petersen, B., Kreyenschmidt, J.: A predictive shelf life model as a tool for the improvement of quality management in pork and poultry chains. Food Control. **29**(2), 451–460 (2013). https://doi.org/10.1016/j.foodcont.2012.05.048

Camacho, D.M., Collins, K.M., Powers, R.K., Costello, J.C., Collins, J.J.: Next-generation machine learning for biological networks. Cell. **173**(7), 1581–1592 (2018). https://doi.org/10.1016/j.cell.2018.05.015

Castelein, B., Geerlings, H., Van Duin, R.: The reefer container market and academic research: a review study. J. Clean. Prod. **256**, 120654 (2020). https://doi.org/10.1016/j.jclepro.2020.120654

da Silveira Petruci, J.F., Fortes, P.R., Kokoric, V., Wilk, A., Raimundo, I.M., Cardoso, A.A., Mizaikoff, B.: Monitoring of hydrogen sulfide via substrate-integrated hollow waveguide mid-infrared sensors in real-time. Analyst. **139**(1), 198–203 (2014). https://doi.org/10.1039/c3an01793a

Digital Container Shipping Association (DCSA): Track & Trace Standards. https://dcsa.org/standards/track-trace/, last accessed 2021/05/28

Defraeye, T., Shrivastava, C., Berry, T., Verboven, P., Onwude, D., Schudel, S., Rossi, R.M.: Digital twins are coming: will we need them in supply chains of fresh horticultural produce? Trends Food Sci. Technol. **109**, 245–258 (2021). https://doi.org/10.1016/j.tifs.2021.01.025

Gambiroža, J. Č., Mastelić, T., Šolić, P., Čagalj, M.: Capacity in LoRaWAN Networks: challenges and opportunities. In: 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech), 1–6 (2019) https://doi.org/10.23919/SpliTech.2019.8783184

Getahun, S., Ambaw, A., Delele, M., Meyer, C.J., Opara, U.L.: Experimental and numerical investigation of airflow inside refrigerated shipping containers. Food Bioprocess Technol. **11**(6), 1164–1176 (2018). https://doi.org/10.1007/s11947-018-2086-5

Gwanpua, S.G., Verboven, P., Leducq, D., Brown, T., Verlinden, B.E., Bekele, E., et al.: The FRISBEE tool, a software for optimising the trade-off between food quality, energy use, and global warming impact of cold chains. J. Food Eng. **148**, 2–12 (2015). https://doi.org/10.1016/j.jfoodeng.2014.06.021

Haass, R., Dittmer, P., Veigt, M., Lütjen, M.: Reducing food losses and carbon emission by using autonomous control – a simulation study of the intelligent container. Int. J. Prod. Econ. **164**, 400–408 (2015). https://doi.org/10.1016/j.ijpe.2014.12.013

Hartgenbusch, N., Borysov, M., Jedermann, R., Lang, W.: Characterization and design evaluation of membrane-based calorimetric MEMS sensors for two-dimensional flow measurement. IEEE Sensors J. **20**(9), 4602–4609 (2020). https://doi.org/10.1109/JSEN.2020.2965332

Janssen, S., Tessmann, T., Lang, W.: High sensitive and selective ethylene measurement by using a large-capacitiy-on-chip preconcentrator device. Sensors Actuators B Chem. **197**, 405–413 (2014). https://doi.org/10.1016/j.snb.2014.02.001

Jedermann, R., Lang, W.: Model based estimation of biological heat generation during cold-chain transport and processing. In: 3rd IIR International Conference on Sustainability and the Cold Chain, St Mary's University, Twickenham, London, UK (2014)

Jedermann, R., Lang, W.: *Computational fluid dynamics modelling of deviating airflow and cooling conditions in banana containers.* In: V International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain – Model-IT 2015, Wageningen, Netherlands, 193-200 (2017) https://doi.org/10.17660/ActaHortic.2017.1154.25

Jedermann, R., Behrens, C., Westphal, D., Lang, W.: Applying autonomous sensor systems in logistics–combining sensor networks, RFIDs and software agents. Sensors Actuators A Phys. **132**(1), 370–375 (2006). https://doi.org/10.1016/j.sna.2006.02.008

Jedermann, R., Behrens, C., Laur, R., Lang, W.: Intelligent containers and sensor networks, approaches to apply autonomous cooperation on systems with limited resources. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics – the Impact on Management, Information and Communication and Material Flow, pp. 365–392. Springer, Berlin (2007a). https://doi.org/10.1007/978-3-540-47450-0_25

Jedermann, R., Gehrke, J.D., Becker, M., Behrens, C., Morales-Kluge, E., Herzog, O., Lang, W.: Transport scenario for the intelligent container. In: Hülsmann, M., Windt, K. (eds.) Understanding Autonomous Cooperation and Control in Logistics – the Impact on Management, Information and Communication and Material Flow, pp. 393–404. Springer, Berlin (2007b). https://doi.org/10.1007/978-3-540-47450-0_24

Jedermann, R., Palafox-Albarrán, J., Barreiro, P., Ruiz-García, L., Robla, J.I., Lang, W.: Interpolation of spatial temperature profiles by sensor networks. Sensors., 2011 IEEE, 778–781 (2011). https://doi.org/10.1109/ICSENS.2011.6127148

Jedermann, R., Nicometo, M., Uysal, I., Lang, W.: Reducing food losses by intelligent food logistics. Phil. Trans. R. Soc. A. **372**(2017), 20130302 (2014a). https://doi.org/10.1098/rsta.2013.0302

Jedermann, R., Lloyd, C., Poetsch, T.: Communication techniques and challenges for wireless food quality monitoring. Phil. Trans. R. Soc. A. **372**(2017), 20130304 (2014b). https://doi.org/10.1098/rsta.2013.0304

Jedermann, R., Praeger, U., Geyer, M., Lang, W.: Remote quality monitoring in the banana chain. Phil. Trans. R. Soc. A. **372**(2017), 20130303 (2014c). https://doi.org/10.1098/rsta.2013.303

Jedermann, R., Praeger, U., Lang, W.: Challenges and opportunities in remote monitoring of perishable products. Food Packag. Shelf Life. **14**(A), 18–25 (2017). https://doi.org/10.1016/j.fpsl.2017.08.006

Jedermann, R., Borysov, M., Hartgenbusch, N., Jaeger, S., Sellwig, M., Lang, W.: Testing Lora for food applications – example application for airflow measurements inside cooled warehouses with apples. Procedia Manuf. **24**, 284–289 (2018). https://doi.org/10.1016/j.promfg.2018.06.026

Jedermann, R., Praeger, U., Geyer, M., Lang, W.: Temperature deviations during transport as a cause for food losses. In: Yahia, E. (ed.) Preventing Food Losses and Waste to Achieve Food Security and Sustainability. Burleigh Dodds, Sawston (2019)

Lu, Y., Liu, C., Wang, K.I.K., Huang, H., Xu, X.: Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. Robot. Comput. Integr. Manuf. **61**, 101837 (2020). https://doi.org/10.1016/j.rcim.2019.101837

Lucklum, F., Janssen, S., Lang, W., Vellekoop, M.J.: Miniature 3D gas chromatography columns with integrated fluidic connectors using high-resolution stereolithography fabrication. Procedia Eng. **120**, 703–706 (2015). https://doi.org/10.1016/j.proeng.2015.08.761

Lütjen, M., Dittmer, P., Veigt, M.: Quality driven distribution of intelligent containers in cold chain logistics networks. Prod. Eng. **7**(2–3), 291–297 (2013). https://doi.org/10.1007/s11740-012-0433-3

Mack, M., Garbade, M., Hebel, M., Kreyenschmidt, J.: Development of a decision support tool for different meat supply chains. In: 5th International Workshop Cold Chain Management, Bonn, Germany (2013)

MOST Tech Sweden AB: Take full control of your supply chain. https://most.tech/ last accessed 2021/05/28

Moureh, J., Tapsoba, M., Flick, D.: Airflow in a slot-ventilated enclosure partially filled with porous boxes: Part II – measurements and simulations within porous boxes. Comput. Fluids. **38**(2), 206–220 (2009). https://doi.org/10.1016/j.compfluid.2008.02.007

Ndraha, N., Hsiao, H.-I., Vlajic, J., Yang, M.-F., Lin, H.-T.V.: Time-temperature abuse in the food cold chain: review of issues, challenges, and recommendations. Food Control. **89**, 12–21 (2018). https://doi.org/10.1016/j.foodcont.2018.01.027

Nunes, M.C.N., Nicometo, M., Emond, J.P., Melis, R.B., Uysal, I.: Quality improvement in fresh fruit and vegetable logistics: berry logistics field studies. Phil. Trans. R. Soc. A. **372**(2017), 20130307 (2014). https://doi.org/10.1098/rsta.2013.0307

Oliveira, C.C.M.D., Oliveira, D.R.B.D., Silveira Junior, V.: Variability in the shelf life of table grapes from same batch when exposed under different ambient air conditions. Food Sci Technol. **41**, 290–300 (2021). https://doi.org/10.1590/fst.14220

Papireddy Vinayaka, P., Van den Driesche, S., Blank, R., Tahir, M.W., Frodl, M., Lang, W., Vellekoop, M.J.: An impedance-based Mold sensor with on-chip optical reference. Sensors. **16**(10), 1603 (2016). https://doi.org/10.3390/s16101603

Pelletier, W., Brecht, J.K., do Nascimento Nunes, M.C., Emond, J.-P.: Quality of strawberries shipped by truck from California to Florida as influenced by postharvest temperature management practices. HortTechnology. **21**(4), 482–493 (2011)

Pitt, J.I., Hocking, A.D.: Fungi and Food Spoilage, vol. 519. Springer, Boston, MA (2009)

Popa, D., Udrea, F.: Towards integrated mid-infrared gas sensors. Sensors. **19**(9) (2019). https://doi.org/10.3390/s19092076

Shashi, Centobelli, P., Cerchione, R., Ertz, M.: Food cold chain management: what we know and what we deserve. Int. J. Supply Chain Manag. **26**(1), 102–135 (2020). https://doi.org/10.1108/SCM-12-2019-0452

Sklorz, A., Janssen, S., Lang, W.: Application of a miniaturised packed gas chromatography column and a SnO2 gas detector for analysis of low molecular weight hydrocarbons with focus on ethylene detection. Sensors Actuators B Chem. **180**, 43–49 (2012a). https://doi.org/10.1016/j.snb.2011.12.110

Sklorz, A., Janßen, S., Lang, W.: Detection limit improvement for NDIR ethylene gas detectors using passive approaches. Sensors Actuators B Chem. **175**, 246–254 (2012b). https://doi.org/10.1016/j.snb.2012.09.085

Stanaway, M.A., Zalucki, M.P., Gillespie, P.S., Rodriguez, C.M., Maynard, G.V.: Pest risk assessment of insects in sea cargo containers. Aust. J. Entomol. **40**(2), 180–192 (2001). https://doi.org/10.1046/j.1440-6055.2001.00215.x

Tahir, M.W., Zaidi, N.A., Rao, A.A., Blank, R., Vellekoop, M.J., Lang, W.: A fungus spores dataset and a convolutional neural network based approach for fungus detection. IEEE Trans. Nanobioscience. **17**(3), 281–290 (2018). https://doi.org/10.1109/TNB.2018.2839585

Tamplin, M.L.: Integrating predictive models and sensors to manage food stability in supply chains. Food Microbiol. **75**, 90–94 (2018). https://doi.org/10.1016/j.fm.2017.12.001

Zaidi, N.A., Tahir, M.W., Vellekoop, M., Lang, W.: Using Allan variance to determine the resolution of ethylene gas chromatographic system. Sensors., 2017 IEEE, 1–3 (2017). https://doi.org/10.1109/ICSENS.2017.8233919

Zaidi, N.A., Tahir, M.W., Vellekoop, M.J., Lang, W.: Design of Novel Ceramic Preconcentrator and Integration in gas chromatographic system for detection of ethylene gas from ripening bananas. Sensors. **18**(8), 2589 (2018). https://doi.org/10.3390/s18082589

Zarkani, S., Rasmussen, C. H.: Remote reefer monitoring looking back and looking forward. In: Cool Logistics Global, 8th global conference, Bremen, Germany (2016)

The Rise of Ultra Large Container Vessels: Implications for Seaport Systems and Environmental Considerations

**Hendrik Jungen, Patrick Specht, Jakob Ovens,
and Burkhard Lemper**

Abstract The development of the global container fleet has followed a clear trend towards ever larger ships over the last 25 years. Particularly striking in this regard is the rise of the dimensionally largest ships, the so-called Ultra Large Container Vessels or ULCVs that can no longer pass through the new locks of the Panama Canal. While recent events such as the six-day blockade of the Suez Canal by the Ever Given have revealed environmental and safety risks of deploying these vessels, even the scheduled and smooth operation generates a whole range of challenges, impacts and costs that come at the expense of external stakeholders. The article aims at identifying these external effects as related to seaport systems as well as environmental considerations by consolidating insights from the scientific and professional discourse.

1 Introduction

Over the past 25 years, the slot capacity of the global container ship fleet has increased continuously as a consequence of a steadily growing number of vessels that have become larger and larger on average. Especially in the years prior to the global financial and economic crisis of 2008/2009, annual increases in total capacity were almost continuously above 10%. During this time, the number of container ships also increased sharply, at times at double-digit annual rates. Since 2010, it can be observed that the number of container ships is only increasing rather slowly, while the average ship size in terms of nominal capacity continues to rise

---

H. Jungen · P. Specht · J. Ovens
Institute of Shipping Economics and Logistics, Bremen, Germany
e-mail: jungen@isl.org; specht@isl.org; ovens@isl.org

B. Lemper (✉)
Institute of Shipping Economics and Logistics, Bremen, Germany

City University of Applied Sciences Bremen, Bremen, Germany
e-mail: lemper@isl.org

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_12

almost linearly. Particularly striking in this regard is the rise of the dimensionally largest ships, the so-called Ultra Large Container Vessels or ULCVs. While this size segment consisted of only eight Maersk E-class vessels with a total capacity of approximately 143,000 TEU before 2012, more than 160 ULCVs with a total capacity of over 3.2 million TEU were already in service at the beginning of April 2021 (CRSL 2021b).

By using ever larger vessels, liner shipping companies are taking advantage of the so-called economies of scale, which have enabled to continuously reduce costs per TEU transported and have given the ULCV its raison d'être (Ge et al. 2019). Any negative effects are still dealt with rather peripherally, and so it is primarily the cost effects of operating ULCVs from an isolated shipping company perspective that dominate the discussion. The fact that the operation of ULCVs also involves costs and impacts on 'external' systems is often overlooked. Critical events such as the accidents of the CSCL Indian Ocean in 2016, the CSCL Jupiter in 2017, the ONE Milano Bridge in 2020 or the current six-day blockade of the Suez Canal by Ever Given clearly show that this is necessary. However, it is not only exceptional circumstances, such as accidents caused by human or technical failures that have a negative impact on areas outside the liner shipping companies. Even the scheduled and smooth operation of such vessels generates a whole range of challenges, impacts and costs that come at the expense of other stakeholders or the general public, often referred to as diseconomies of scale (Stopford 2009). Likewise, there are also positive and often misappropriated effects of ULCVs that are not directly related to operational profitability.

This article aims at identifying these external effects as related to seaport systems as well as environmental considerations by consolidating insights from the scientific and professional discourse. In order to illustrate the future importance of ULCV within the shipping market, the development of the container ship fleet over the last 25 years is first outlined and the increasing capacity share of the largest units elaborated afterwards. Since there currently exists no uniform understanding of which determinants make a large container ship a ULCV, Chapter "Autonomous Control of Logistics Processes: A Retrospective" of this article provides a clear definition. Furthermore, estimates on future vessel size developments will be presented as well. Chapter "Explorable Uncertainty Meets Decision-Making in Logistics" deals with the implications that increasing vessel sizes have for seaport systems. This includes not only necessary adaptions of port infra- and superstructures but also operational performance requirements that shipping lines impose on terminal operators. As operational constraints of ULCVs influence port choice decisions of shipping companies, ongoing and future effects on port network structures are discussed as well. In Chapter "Complex Networks in Manufacturing and Logistics: A Retrospect" the most relevant environmental impacts of ULCVs are outlined. After an outline of the different impacts is given, possible benefits that ULCVs have over smaller vessel units are considered. Subsequently, the improvements of ULCVs energy efficiency are discussed in the context of emission regulations. The section concludes with the environmental impacts of dredging operations, which are conducted to accommodate ever larger vessels.

2 The Rise of ULCVs

2.1 Towards Gigantism and Segmentation in Container Shipping

While the number of ships within the global container fleet has increased marginally by 9% since 2011 (from 4933 in April 2011 to 5450 units in April 2021), capacity in terms of container slots has grown by more than 65% (from 14.4 to 23.8 million TEU) (see Fig. 1). This relatively high capacity growth compared to the small number of new buildings indicates that container ships are becoming larger on average. In fact, the average nominal capacity of a container vessel increased by a total of 185% between 1996 and 2021, respectively, from 1524 to 4351 TEU per ship (CRSL 2021b). The strongest driver of this constant growth in size were and are the economies of scale that ultimately led shipping lines with larger units to the deployment of ULCVs (Ge et al. 2019). While economies of scale are very distinct for comparably smaller units, however, the positive effect diminishes with increasing size and so the question to which ship size further capacity growth is still worthwhile has been the subject of controversial debate for many years (Ge et al. 2019). Stopford, for example, argues that economies of scale below 4000 TEU are more evident than above (Stopford 2009). With 7500 TEU Sys gives a somewhat higher threshold than Stopford, but also states that economies of scale for the shipping lines decrease sharply as units become larger (Sys et al. 2008). Recent studies show that under favourable market conditions, with a slot utilization of over 90% and high freight rates, 25,000 TEU vessels can still achieve significant advantages over smaller units with capacities of 18,000 or 20,000 TEU (Ge et al. 2019). It is therefore hardly surprising that order activities for container ships with

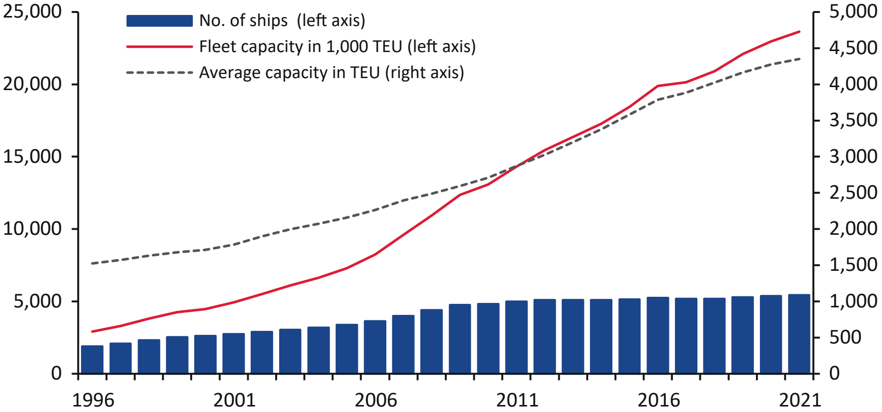
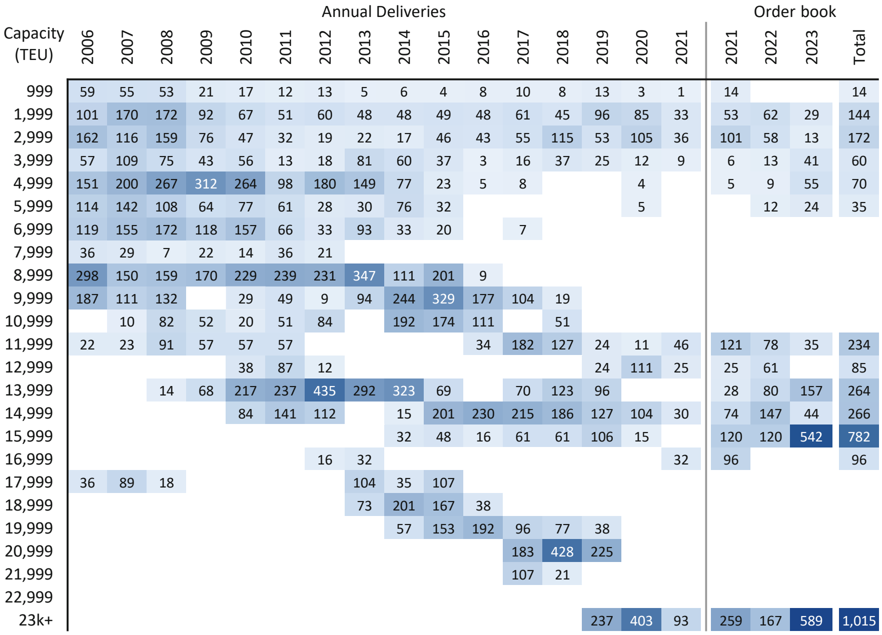


Fig. 1 Development of world container fleet according to the number of ships, total nominal fleet capacity and average nominal capacity per ship (CRSL 2021b)



Capacity (TEU)	Annual Deliveries																Order book			
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2021	2022	2023	Total
999	59	55	53	21	17	12	13	5	6	4	8	10	8	13	3	1	14			14
1,999	101	170	172	92	67	51	60	48	48	49	48	61	45	96	85	33	53	62	29	144
2,999	162	116	159	76	47	32	19	22	17	46	43	55	115	53	105	36	101	58	13	172
3,999	57	109	75	43	56	13	18	81	60	37	3	16	37	25	12	9	6	13	41	60
4,999	151	200	267	312	264	98	180	149	77	23	5	8			4		5	9	55	70
5,999	114	142	108	64	77	61	28	30	76	32					5			12	24	35
6,999	119	155	172	118	157	66	33	93	33	20		7								
7,999	36	29	7	22	14	36	21													
8,999	298	150	159	170	229	239	231	347	111	201	9									
9,999	187	111	132		29	49	9	94	244	329	177	104	19							
10,999		10	82	52	20	51	84		192	174	111		51							
11,999	22	23	91	57	57	57					34	182	127	24	11	46	121	78	35	234
12,999					38	87	12							24	111	25	25	61		85
13,999			14	68	217	237	435	292	323	69		70	123	96			28	80	157	264
14,999					84	141	112		15	201	230	215	186	127	104	30	74	147	44	266
15,999									32	48	16	61	61	106	15		120	120	542	782
16,999							16	32								32	96			96
17,999	36	89	18					104	35	107										
18,999								73	201	167	38									
19,999									57	153	192	96	77	38						
20,999												183	428	225						
21,999												107	21							
22,999																				
23k+														237	403	93	259	167	589	1,015

Fig. 2 Containership deliveries according to nominal capacity as of 04th April 2021 (in 1000 TEU) (CRSL 2021b)

capacities of more than 23,000 TEU have picked up strongly since the fourth quarter of 2020, as charter rates for container ships have only known one direction since the multi-year low in June 2020, and that is up (CRSL 2021a). As of the beginning of April 2021, out of the total of around 589,000 TEU in the order book in this size-class, which are to be delivered in 2023 (see Fig. 2), nearly 474,000 TEU have been ordered since October 2020 (CRSL 2021a).

Until around 2011, the nominal total capacity of the global container ship fleet was distributed rather homogeneously among various ship sizes up to a maximum of 14,999 TEU, with the deliveries of Maersk E-class ships in the years 2006 to 2008 standing out clearly (see Fig. 2). Since 2012, a segmentation into merely three size segments is observable. The first segment consists of smaller units up to 5999 TEU, with Feeder- and Feeder-Max-vessels making up the major share. The second segment includes large units with nominal capacities between 11,000 and over 16,000 TEU, which are still able to pass the Neo-Panamax locks due to their dimensions. In this segment, it can be observed, in particular with regard to the units in the current order book, that ships over 15,000 TEU represent the largest subgroup in terms of capacity. Since 2020, the third segment has consisted exclusively of vessels with a nominal capacity of over 23,000 TEU. This means that since 2020, no vessel with a nominal capacity between 17,000 and 22,999 TEU has been delivered or ordered, so that it can be said with a fair degree of certainty that future fleet growth will take place exclusively above 23,000 TEU. If all units in the order book are delivered as planned by 2023, this size segment will be more than double by then (CRSL 2021b).

2.2 *Too Big for the Panama Canal*

The term Ultra Large Container Vessel (ULCV) emerged in the early 2000s to describe model designs that were beyond the largest container ships of the time in terms of dimensions and capacity. At that time, when the so-called Maersk C-class (LOA[1] 345 m, Beam 42.8 m, Draught 14.5 m, Capacity 8650 TEU) was the largest container ship ever built (CRSL 2021a), there were concepts for ships with up to 12,500 TEU where the term ULCV or Ultra Large Container Ship (ULCS) was already in use (van Ham 2005, p. 89). Albeit ships with more than 18,000 TEU seemed feasible from a shipbuilding point of view (Wijnolst et al. 1999, p. 74), it was assumed that the limitations, in particular port limitations, associated with their size would prevent widespread acceptance of such vessels (van Ham 2005). The units of up to 24,000 TEU that are in service today falsified this assumption.

To date, there is no clear and uniform definition of when a container ship is considered a ULCV and the term is usually used for the largest existing ships at the time. In general, it is common practice to divide container ships into certain classes based on their nominal slot capacity in TEU. The various relevant thresholds for ULCVs are, for example, 14,500 TEU (Sahoo 2021), 15,000 TEU (Lian et al. 2019) or even 18,000 TEU (Heaney et al. 2020). In some cases, the size restrictions of the new locks of the Panama Canal are cited as relevant differentiator for vessels in addition to the nominal capacity (Heaney et al. 2020; Sahoo 2021). Additionally, the terms MGX-23 and MGX-24 (pronounced Megamax) are used for ships with a width of 23 and 24 container rows, respectively, with nominal capacities of around 18,000 to almost 24,000 TEU (Alphaliner 2017). Because of this multitude of different definitions, it is important to accurately define the term ULCV for the purposes of this article.

Looking at the growth in ship size over the past decades (see Fig. 3), it is noticeable that the increases of the nominal capacity between 1996 and 2005 were almost exclusively achieved by longer ships with a slightly increased draught, while the beam stagnated at 42.8 m. The commissioning of the Maersk E-class in 2006 opened a new chapter in ship size growth. With a LOA of almost 400 m, a beam of over 56 m and a draught of up to 15.5 m, these ships were significantly larger than container ships built up to that time. The capacity is given by the classification society American Bureau of Shipping as 11,000 TEU (American Bureau of Shipping 2021), while other sources state the capacity as 15,000 (Haas 2016). This discrepancy can be explained by the use of different calculation methods

---

[1] 'Length over all', which describes the maximum length of a vessel's hull.

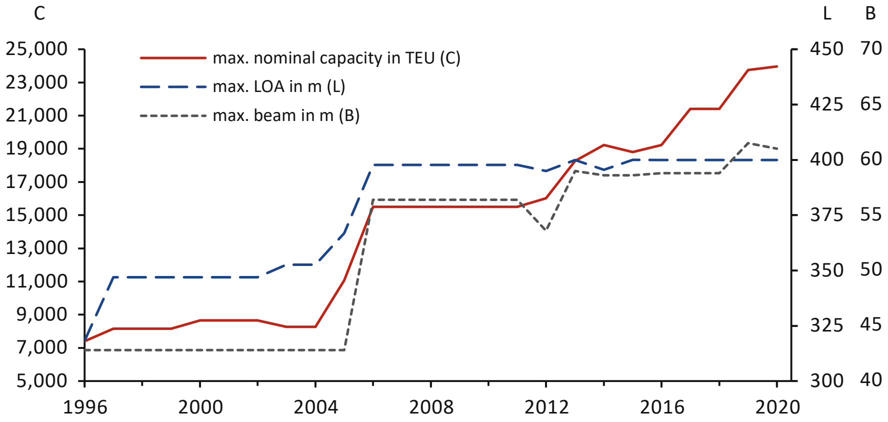


Fig. 3 Development of container ship sizes by nominal capacity, LOA and Beam, based on the largest ship in the respective year by nominal capacity (CRSL 2021b)

to determine the capacity. While the classification society has assumed a pre-specified weight/TEU condition (here 14 tonnes/TEU), a so-called homogeneous loading condition, the other values describe the nominal capacity, that means the maximum allowable number of containers, in terms of SOLAS navigational bridge visibility requirements (Koutroukis et al. 2013; Haas 2016; IMO 2020b). By increasing the height of the deckhouse and the lashing bridges, the nominal capacity of the Maersk E-class was increased to 17,816 TEU in 2018 (THB 2018; CRSL 2021b). The Maersk E-class was replaced by CMA CGMs Explorer-class as the largest container ship in terms of nominal capacity in 2012. An example for one of the vessels of this class is the CMA CGM Marco Polo. While the Explorer-class is shorter and narrower than the Maersk E-class, the maximum draught has been increased to 16 metres. In particular, the new twin-island design, i.e. the location of the bridge at the foreship section with the engine room located at semi-aft, resulted in improved visibility and has significantly increased nominal capacity. When applying the homogeneous calculation method with 14 tonnes/TEU, vessel operator CMA CGM indicates the capacity with 11,358 TEU, what is more or less equal to the capacity of the Maersk E-class.

The next significant step since the Maersk E-class towards ships, which were actually and not only capacity-theoretically larger, was taken in 2013 with the Maersk Triple E-class. With a length of still approximately 400 m and a beam of 59 m these twin-island ships can accommodate 23 container rows and thus are a whole container row wider than the Maersk E-class (CRSL 2021b). With these ships, the era of slow steaming has also entered the design specs of the ULCV. A comparison of the Triple E-class' engine performance with that of the Maersk A-class from 2003 shows that it is almost identical, with nearly twice the deadweight (CRSL 2021b). Because of their 23rd container row, the Maersk Triple E-class and other ships of comparable size are also called MGX-23 (Alphaliner 2017). From 2013 up to and including 2018, the dimensions of the largest container ships have remained virtually the same. While new nominal capacity records have been set every year, these are merely theoretical and range between 18,000 and 21,000 TEU. In 2017, the term MGX-24 emerged to refer to a new generation of containerships with an additional 24th container row that entered service in 2019 (Alphaliner 2017). While still around 400 m in length, these ships are more than 61 m wide, have a draught of up to 16.5 m and a nominal capacity of more than 23,000 TEU (CRSL 2021b).

Based on the above, it has been shown that looking at nominal capacity data as a sole determinant of ship size can be misleading. Instead, static parameters, i.e. length and beam of ships, must be included in the consideration. As presented in the following sections, the draught has significant effects on the seaport system but as a dynamic parameter it is not suitable for grouping purposes. In this article, ULCVs are defined as container ships that can no longer pass through the new locks of the Panama Canal due to their static dimensions of LOA and beam. This means that container ships with a LOA greater than 367.28 m and/or a beam greater than 51.25 m (Panama Canal Authority 2021) are regarded as ULCVs. Within the segment, in particular the increased beams for 23 container rows from 2013 and for 24 container rows from 2019 onwards represent significant development steps.

However, since almost all statistics on the container ship fleet are subdivided according to nominal capacity, an attempt is made to establish comparability here. In terms of capacity, currently almost all ships with a nominal capacity of over 15,000 TEU fall into the just mentioned definition of ULCVs in terms of dimensions. Of the 184 units in service with more than 15,000 TEU, only seven are still able to transit the Neo-Panamax locks of the Panama Canal (CRSL 2021b). These seven vessels entered service in 2019 and 2020 and have a nominal capacity between 15,052 and 16,010 TEU (CRSL 2021b). This again demonstrates the difficulty of assigning ship size classes based on nominal capacity. While all vessels with a nominal capacity above 15,000 TEU that were built until 2018 have dimensions that make transiting through the Panama Canal impossible, there are newer designs that push the theoretical capacity limit further up. The aforementioned 16,010 TEU nominal capacity currently represents the new capacity limit to Neo-Post-Panamax vessels and thus to ULCVs in the sense of this article. Ergo, while a TEU threshold of 15,000 TEU seems reasonable for the consideration of the fleet development until 2018, this threshold would have to be set higher from 2019 onwards. With the exception of three ships, all ships built after 2019 that no longer fit through the Panama Canal due to their dimensions have a nominal capacity of at least 19,000 TEU. In addition, the current order book does not show a single order in the capacity range between 16,010 and 23,000 TEU. Thus, from 2019 onwards, a capacity limit of 19,000 TEU would cover all ULCVs in the sense of this article, except three ships of the Maersk H-class with a nominal capacity of 15.226 TEU that cannot transit the Neo-Panamax locks due to their beam (CRSL 2021b). All ULCVs currently in service have been deployed on the main east-west trades between Europe and Far East (CRSL 2021b), and pushed medium-sized ships into other trade routes, where these ships in turn displace even smaller units. This phenomenon is known as the cascading effect (Cariou and Cheaitou 2014).

2.3 Where Do we Grow from Here: Bigger Vessels Yet to Come?

An answer to the question of where the growth in size in container shipping will lead can only be speculative at present. Malchow considers that construction of the first 30,000 TEU ships (nominal capacity) can start as early as 2025. Malchow bases his forecast on the exponential continuation of the development of the nominal slot capacity of the respective largest ships in terms of capacity since the end of the 1960s (Malchow 2017). The dimensions of these 30,000 TEU container ships are the subject of controversial debate. While Park et al. give the dimensions of these ships, determined by a regression analysis, as 453 m LOA, 72.0 m beam and up to 17.3 m draught (Park and Suh 2019), new forecasts are much more modest here. Thus, according to a recent publication by Alphaliner, the 30,000 TEU mark could be reached with dimensions of 425 m × 66.1 m, which would involve an extension of two 40' bays compared to the MGX-24 ships and a widening to 26 rows. Until then, appropriate intermediate steps would be taken over length and width, for example by stretching the vessel to 425 m while adding only one extra row to 25 rows in total (Alphaliner 2021).

3 Implications for Seaport Systems

3.1 Necessity to Adapt to ULCV Requirements

As the example of the Panama Canal illustrates, the relationship between bottlenecks onshore and the developments in vessel sizes is one of reciprocal influences. This also holds true for seaport systems that provide the link between maritime and hinterland transportation (Carbone and Martino 2003) and thus ensure a transformation of mass to single transports (and vice versa) (Schönknecht 2009). While it can be argued that the 'gigantism in containerships' has been significantly stimulated by efficiency gains in container terminals over the last decades (Haralambides 2019), the rise of ULCVs has induced a number of challenges that seaports currently face. As will be shown in the following section, those challenges are not just apparent due to the sheer dimensions of those vessels but also due to the performance requirements that are connected with executing high-intensity port calls at maximum pace. Ultimately, the capabilities of port systems to adapt to current and future vessel sizes will most likely affect operational choices by shipping alliances and thus will have an impact on the structure of future port and transport networks.

3.2 Maritime Accessibility No Longer a Matter of Course

The dimensions of container vessels have taken a steep curve in the last 25 years. By making heavy investments, most ports have managed to adapt to these developments and thereby maintaining their ability to accommodate vessels nautically as well as serving them operationally at berth (Sys et al. 2008). However, in light of the recent and future generations of ULCVs, maritime accessibility is not a matter of course any longer and often not given at all or only under certain conditions (Notteboom 2016). From a nautical point of view, access to the port is constrained by the dimensions of the approach channel, which links 'the berths of a port and the open sea' (PIANC 2014) as well as turning basins and berths within the port. To safely navigate a vessel, sufficient space, air draught[2] as well as under keel clearance needs to be guaranteed at all times (PIANC 2014). As outlined in the previous section, today's ULCVs come with a maximum draught of 16.5 m, while the individual draught depends on the load situation of the vessel. A closer look on draught restrictions of selected container ports indicates that fully loaded ULCVs often already reach or exceed the set limits (Sys et al. 2008; Rodrigue 2021). Well-known examples include the Northern European ports of Antwerp, Bremerhaven and Hamburg, which are currently restricted to draughts of less than 16 m even at high tide (Hansestadt Bremisches Hafenamt 2021; Scheldecoördinatiecentrum Vlissingen 2021; Hamburg Port Authority 2018). The same situation can be observed in North America, e.g. in ports located at the Gulf of Mexico and several East Coast ports in the USA (Rodrigue 2021). As a result, water depth has become a major bottleneck for many ports in their quest to accept the largest container vessels.

By making use of tidal windows and specific speed and weather limits (Vantorre et al. 2014), some port and traffic authorities follow 'access policies' that allow for accommodation for ULCVs under strict conditions. Also, shipping companies may decide to reduce vessel draughts, e.g. by changing the port rotation sequence so that ports with restrictions will be called at a lower load factor (Merk et al. 2015; Det Norske Veritas AS 2013). However, all these organizational measures come with less flexibility and increased costs for shipping companies and thus make the transport less profitable (Sys et al. 2008). A common approach to improve the water depth conditions is to deepen and widen berthing areas, turning basins and approach channels through dredging. It involves 'the operation of removing material from one part of the water environment and relocating it to another' (European Dredging Association n.d.). However, in addition to high costs, physical and environmental limitations need to be considered in regard to dredging operations. The problem intensifies for 'inside ports' that face even more unfavourable conditions with respect to available space, natural water depths and thus costs and environmental impacts (Andrés and Piniella 2017). As various cases of the near past have

---

[2] Air draught needs to be taken into account with respect to crossing bridges. Prominent examples are the Koehlbrand Bridge in Hamburg or the Tsing Ma Bridge in Hong Kong, which do not allow for enough air draught for ULCV even at favourable tidal conditions.

illustrated, the increased dimensions of ULCVs ultimately make navigating and manoeuvring more difficult and allow for significantly less margin of error (Mallin and To 2016).

Further conditions to handle vessels alongside the berth include adaptions of quay structures as well as superstructure to carry out loading and unloading processes. This comprises quay structures and mooring systems that are dimensioned sufficiently to handle increased bollard pull forces of larger vessels (Mallin and To 2016; Stoschek et al. 2018; Soderberg 2017). Not just the added forces induced by the vessel itself put additional strain on quay structures, but also the weight and leverage of ship-to-shore (STS) cranes that allow processing all tiers and rows of the vessel (Park and Suh 2019). As shown in Sect. 2.2, vessels that fit with our definition of ULCVs come with a beam of 59–61.5 m and can carry 23–24 rows across and 10–11 tiers on deck. Thus, ULCVs induce special requirements on the cranes outreach[3] as well as on the maximum lifting height (Park and Suh 2019). Cranes that meet these demands are often referred to as 'Super Post-Panamax' (Čerin and Beškovnik 2020) or 'Megamax' cranes (Liebherr Group 2021). The largest STS cranes to date come with an outreach of 80 m, a lifting height of 54 m and allow for handling of 26 rows across (Hill 2020a, b). Thus, it can be expected that these cranes will be able to process future ULCV generations (see 2.3).

As shown in Fig. 4, the worldwide STS crane deliveries of the near past reflect the growth of vessel sizes. With respect to outreach for example, the share of new cranes that are able to serve 23+ rows and come with an outreach of 60.0 m and above went from of 38% in 2011 to 81% in 2017. The same trends can be observed for other parameters such as lifting height and lifting weight (PEMA 2018). The developments on the equipment market indicate that terminal operators are actually trying to cope with and prepare for increased vessel sizes by adapting their superstructure.

3.3 Increased Pressure to Squeeze Berth Productivity

Assuming a slot utilization similar to smaller vessels, it becomes obvious that increased container capacities on ULCVs are likely to lead to a higher number of containers to be handled per port call, commonly referred to as the 'call size'. Due to this, practical observations from experts as well as empirical studies show that ULCVs experience significantly longer port stay times compared to smaller units (Specht et al. 2020; JOC Group 2014).

---

[3] The outreach needs to account for the beam of the vessel as well as the distance from the vessels hull to the centre of the cranes' rail system at berth. As a 'rule of thumb', the necessary outreach in meter can be estimated by multiplying the number of rows by 2.5 m and adding 5 m for the distance between the crane and the vessel (PEMA (2018)). Hence, a vessel with 24 rows would require an outreach of more than 65.0 m.

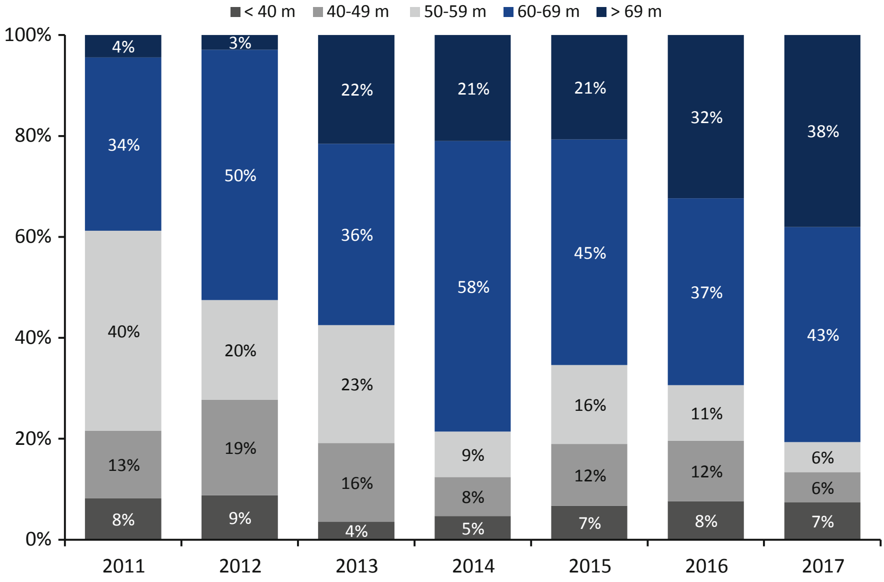


Fig. 4 Global STS crane deliveries by outreach (2011–2017) (PEMA 2018)

Due to increased capital intensity of ULCVs paired with slow steaming strategies, non-productive port stay time endangers profitability of those services (Haralambides 2019; Lane and Moret 2014). As a result, carriers exert enormous pressure on terminal operators to increase handling efficiency and consequently reduce the duration of port calls. The predominant metric that allows for comparisons between ports and terminals is ‘gross berth productivity’, which compares the number of loading and unloading operations per call with the total port stay time (JOC Group 2014). In general there are two main components affecting the ‘operational’ part of berth productivity, being the speed and productivity of the cranes, usually quantified as moves per hour, and the number of cranes deployed on the vessel, also referred to as ‘crane intensity’ (Merk et al. 2015).

Even though increased outreach and lifting heights result in higher travel distances, it can be observed that crane manufacturers have managed to increase cycle speeds of Megamax cranes compared to older and smaller ones. Among other technological solutions, Megamax cranes usually allow to use twin-lift or even tandem-lift spreaders, which enable to transport two containers at the same time, i.e. up to 4 TEU in one move (Böse 2020). Besides the physical capabilities of the crane itself, the efficiency of quay-to-yard transportation systems, which need to guarantee a seamless flow of containers from and to the crane, is a prerequisite to maximize cycle speeds (Lane and Moret 2014). Crane density as the other setscrew needs to account for blocking and safety constraints between adjacent cranes as well as the layout of the respective ship (e.g. position of bays and bridges) and is therefore constrained for each type of vessel (Martin et al. 2015). Furthermore, different operational factors such as the availability of cranes and workforce, crane occupation on other vessels at berth or contractual commitments influence how many cranes can and will be deployed during a specific port call. As a result, it can be observed that even though a 400 m vessel could theoretically be worked on with more than 12 cranes simultaneously (Martin et al. 2015), most ports do not deploy more than six (The loadstar 2015) to eight (Martin et al. 2015) cranes on one ULCV.

Given these conditions, the performance requirements set by shipping companies are rather challenging. As an example, the carrier 'Maersk' has stated a target of up to 5000–6000 moves per day which corresponds to 208–250 moves per hour for ULCV calls (WorldCargo News 2011). Despite the advancements in equipment technology, ports still struggle to meet targets. A report from JOC Group shows that in 2013, the berth productivity range of the ten best performing terminals worldwide ranged from 131 to 179 moves per hour for vessel sizes larger than 8000 TEU. Clear differences can be observed with respect to the different regions worldwide. European and American terminals significantly lag behind those from Asia in this context (JOC Group 2014).

From the terminals' perspective, an often-reported obstacle is increasing ULCV delays that prevent terminals from making optimal planning decisions (Sys et al. 2008). In addition, terminal operators report poorly stowed vessels that necessitate unproductive shuffle moves and thus lead crane optimization efforts to fail (Haralambides 2019). Further, the largest container vessels are no longer increasing in length but in beam and height, which does not allow for higher crane intensity (Martin et al. 2015). Longer port stays and intensified nautical constraints are also observed to lead to 'knock-on effects' (Lane and Moret 2014) that ultimately result in further off-schedule ULCV calls. Due to the high capacity and operational requirements (e.g. berthing space and crane capacity), it is very difficult for terminals to ensure flexibility or even to make up for delays. Better coordination between shipping companies and terminals is therefore vital to improve this situation (Lane and Moret 2014).

3.4 Coping with Peaks Induced by ULCVs

Large call sizes in combination with increased berth productivity ultimately lead to an intensified workload that is spread over a short period of time (Martin et al. 2015; Lane and Moret 2014), furthermore referred to as 'peaks'. As we have outlined in the previous section, quayside operations directly link to other terminal subsystems, such as the yard storage area and the landside interface (Čerin and Beškovnik 2020).

Since containers to be loaded have to be provided in due time at berth and unloaded ones to be transported to the terminal yard, storage operations need to be efficient and scaled accordingly in order not to constrain quayside operations (Lane and Moret 2014). It can be observed that ULCV induced peaks correlate with higher storage utilization in terminal yards (Merk et al. 2015). As storage capacity cannot be extended endlessly, terminal operators try to make better use of available storage areas by reducing container dwell time through economic incentives for port users e.g., by increasing storage costs or reducing free storage time (Čerin and Beškovnik 2020).

Peaks also affect the landside interfaces of terminals, as incoming and outgoing containers need to be delivered and picked up in temporal relation to the port call. In fact, observations indicate that there exists a relationship between call sizes and increased hinterland traffic volume, especially on the road (Merk et al. 2015). Depending on the individual circumstances, an extension of port-related hinterland infrastructure may be necessary, as traffic volumes can exceed existing capacities (Russell et al. 2020; Mongelluzzo 2020). Especially ports with a high modal share of road transportation show increased gate congestion in relation to arrivals of larger vessels (Ozbas et al. 2014). Some terminals try to alleviate this issue through optimized gate processes such as the introduction of truck appointment systems that are meant to ease traffic volumes and make use of off-peak times (Merk et al. 2015).

The uneven distribution of workloads within the different subsystems of the terminal makes it difficult to provide the right amount of capacity while obtaining own economic objectives. This not only applies for superstructure and other equipment, but also to the available workforce. In this way, ULCV induced peaks reverse the trend of regularized and plannable personnel deployment, which has once been achieved through the standardized processes of modern terminals (Merk et al. 2015). Furthermore, peak effects do not just touch the physical flow of cargo but also impact administrative processes, data flows or custom procedures which need to be dimensioned properly as well (Merk et al. 2015). It can be expected that terminals with a high degree of automation have a better position to flexibly cope with ULCV induced peak workloads (Rijsenbrij and Wieschemann 2020).

3.5 Remodelling Seaport Network Structures

In their quest to optimize operations for ULCVs, shipping companies will take decisions that ultimately have great influence on port network structures (Sys et al. 2008). Thanks to increased bargaining power due to the consolidation within the shipping market, carriers persistently pursue to push berth productivity improvements in order to minimize non-productive port stay time. However, terminal operators have a hard time meeting these demands. An additional challenge is maritime access, which is either not given at all or only under certain constraints that are likely to limit flexibility and increase costs (see 3.2).

In addition to the economic constraints associated with the realization of economies of scale, the aforementioned aspects lead to the fact that shipping companies are reducing the number of port calls on ULCV services to fewer hubs, which leads to the emergence of hub-and-spoke networks (Haralambides 2019). Within these networks, port choice for ULCVs will especially be based on maritime accessibility, operational performance (Musso and Sciomachen 2020), connectivity but also on strategical considerations (Ducruet and Notteboom 2012. ULCV-ready ports within important trade areas that serve almost identical hinterlands find themselves in strong competition to stay one of the network hubs (Sys et al. 2008). Within these network configurations, carriers will constantly need to trade off higher slot utilization of their ultra large vessels against additional handling costs that occur due to necessary transhipment. Hence, it can be suggested that ‘hub ports’ with strong local hinterlands that can maintain a high share of import/export cargo will have the best conditions to maintain a sustainably important role within liner service networks (Ducruet and Notteboom 2012).

Experts suggest that maritime accessibility is likely to become an even more important factor of port choice for ULCVs. As a consequence, non-urban ports could take on an increasingly important role compared to ports in urban areas that do not provide enough room for adaptations (Merk 2018). An example provides the currently ongoing debate on Hapag-Lloyd’s recent announcement to reallocate ULCV services from the urban port of Hamburg to the greenfield terminal in Wilhelmshaven (WorldCargo News 2021a). Calls for stronger environmental restrictions as well as the risk of casualties in vulnerable urban areas add to these discussions.

4 Environmental Considerations of ULCVs

4.1 Ecologies of Scale

The impacts that the global shipping industry has on the environment are manifold. To encounter negative consequences, international regulations have been set in place that are often amended by national policies. In the following, an overview of the major possible impacts, shipping has on the environment, will be given with the affiliated regulations that are already in place.

Various studies have shown that certain components of anti-fouling paints are highly toxic to the marine environment. As the hull of a ship is almost permanently exposed to seawater, particles of the paint can enter the environment by abrasion or by disintegration. Scientific investigations have proved that these particles would poison marine life forms and enter the entire food chain. Therefore, certain components were banned from the use in anti-fouling paints since 2008 (MEPC 1990; IMO 2001).

At the end of a ship’s lifetime, which means it is either damaged or further operation does not pay off any longer for economic, ecologic or regulatory reasons, it gets wrecked and any valuable parts are recycled. Over the ship’s lifecycle, many potentially hazardous materials accumulate in a ship. These are, for example, residues of fuel oils, lube oils, cargo, paints. MEPC released different resolutions as guidelines to reduce environmental impacts of ship recycling (MEPC 2011a, 2012a, b, d, e, 2015).

The disposal of garbage at sea is another issue that endangers marine life and the marine environment and is regulated by MARPOL Annex V (MEPC 1973), which is updated at regular intervals. Ships above a certain size have to carry a garbage management plan and have to comply with the standards for shipboard incinerators (MEPC 2012c, 2014a, 2017).

Ships are also emitters of a considerable degree of noise. While there is regulation in place to protect the crew and the passengers from shipboard noise, there currently exists no regulation to protect the marine environment from ship noises (ILO 2006; MSC 2012a, b). The issue of ship noise impacts on marine mammals was discussed already in 2007, but currently only non-mandatory guidelines are in place (MEPC 2007, 2014b).

Another environmental issue are the exhaust emissions of ULCVs that are the direct result of the fuel used in a ship's internal combustion engine. Exhaust emissions include greenhouse gases (GHG) and other components. Most notably of these other components are nitrous oxides (NO_x), sulphur oxides (SO_x) and particulate matter (PM) (Faber et al. 2020). The amount and the composition of the exhaust emissions depend on the specific machinery and the fuel in use. On modern day container vessels, different types of machinery are operated on different fuels. Therefore, the operation of vessels leads to a wide range of different emissions that are subject to various regulations that will be discussed detailed in Sect. 4.2.

None of these environmental impacts is unique to ULCVs and happens on any vessel. However, as has been discussed in Sect. 2.1, ULCVs have been developed in recent years with an increased nominal capacity while still maintaining the same dimensions. That means, the environmental impacts, that have been been discussed above are staying roughly the same for each vessel, while they can be attributed to an increasing amount of containers. As a result the environmental impact per TEU is decreasing. This effect could best be described as 'ecologies of scale' (Schlich and Fleissner 2005).[4]

4.2 Regulating Emissions and Improving Energy Efficiency

These 'ecologies of scale' become especially apparent if the exhaust emissions of ULCVs are compared. Besides carbon dioxide (CO_2) a vessel's combustion engines also emit other GHGs that are measured in carbon dioxide equivalents (CO_2e). In 2018 the combined CO_2e-emissions of shipping, which includes international and domestic shipping as well as fishing, contributed to only 2.89% of the yearly global anthropogenic emissions. The shipping sector is dedicated to reduce its

---

[4] The term 'ecologies of scale' according to the author describes the effect that the ecological impact of a production line decreases as the amount of produced items is increased, similar to the term 'economy of scale". In the context of this article, the amount of items translates to a service provided, i.e. the amount of containers transported.

total CO_2-emissions by 50% until 2050 in relation to 2008 levels (Faber et al. 2020; MEPC 2018). Because of a growing global trade volume and consequently a growing number of ships with a growing amount of cargo being underway, this requires a reduction of CO_2e-emissions by 85% per individual ship (IMO n.d.). To tackle this issue, the Marine Environment Protection Committee (MEPC) of the IMO proposed various measures.[5] One is the Energy Efficiency Design Index (EEDI), which evaluates the energy efficiency of a vessel by defining a minimum energy efficiency baseline that new built ships need to achieve and that is increased every 5 years (MEPC 2011b, 2020), to ensure constant improvements of new built vessels. While the EEDI only aims at improving the efficiency of new built ships, MEPC adopted a similar index for the existing fleet, the Energy Efficiency Existing Ship Index (EEXI) in 2021 to be applied after 2022 (IMO 2021). Both indices, however, have been criticized to only make small contributions towards IMO's decarbonization goals (Trivyza et al. 2020; Rutherford et al. 2020; Psaraftis and Kontovas 2021). EEDI and EEXI are accompanied by a so-called Ship Energy Efficiency Management Plan (SEEMP) and the Carbon Intensity Indicator (CII). The SEEMP is an operational mechanism to push ship owners to optimize the operational efficiency of a ship (MEPC 2011b), while the CII is an annual rating of the CO_2-emissions of any singular vessel, which can be tied to incentives for a shipping company to invest in cleaner technologies (IMO 2020a, c, 2021).

The emission of NO_x, SO_x and PM is regulated by MARPOL since 1973. While these emissions do not have a global warming potential such as GHG, they come with negative local effects on populated areas such as cities in proximity to ports or navigation routes. Since the first of January 2020 (MEPC 2016) fuels have a maximum permissible amount of SO_x and PM of 0.5% (MEPC 2019). Especially in ecologically sensitive regions, the so-called Emission Controlled Areas (ECA), ships need to comply with even stricter requirements regarding machinery and fuel. Vessels built after the first of January 2016 that are sailing in an ECA need to be equipped with a Tier-III-engine (MEPC 2018, 1973) and the maximum permissible amount of SO_x and PM in the fuel is limited to 0.1% (MEPC 1973).

Due to these regulations, ULCV designs are constantly being improved with respect to their fuel efficiency. While the regulations are relevant for any ship, they also find application on ULCVs. However due to the 'ecologies of scale' ULCVs provide ecological benefits over smaller vessels and are constantly improving. In this section, an analysis of the energy efficiency changes of ULCVs over time is conducted, to estimate the improvements that have been made. For every year the fuel consumption of the largest ship by nominal capacity that has been delivered in that respective year has been estimated according to the formulas of the IMO (Faber et al. 2020). To calculate the fuel consumption, it was assumed that the vessels are

---

[5] Also, the European Union is pushing the maritime sector to reduce the fuel consumption its CO_2-emissions. In 2015 the EU adopted legislation to a program to monitor, report and verify CO_2-emissions from the shipping sector that has been proposed in 2013. The aim of the EU is to take regulations to a practical level and use the findings in the discussions with the IMO (European Commission (2013, 2020)).

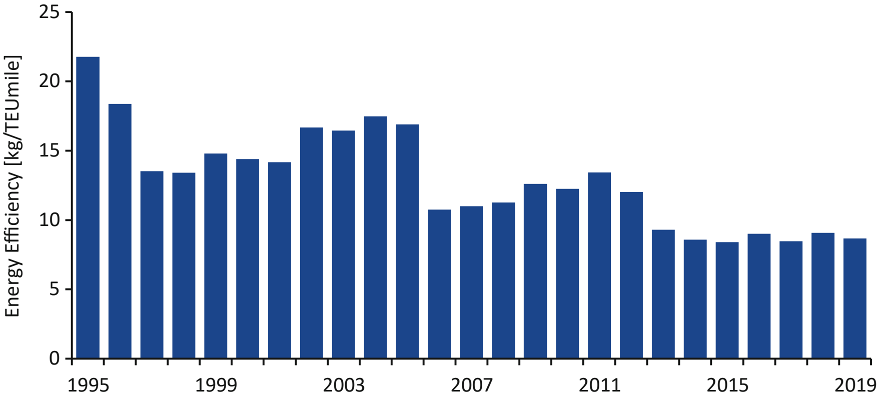


Fig. 5 Development of Energy Efficiency of ULCVs. Own diagram based on CRSL (2021a)

operated at a uniform speed to compare them from a current point of view. Only the fuel consumption of the main engine is part of the estimation. The calculation does not account for any consumption from auxiliary engines or boilers, the amount of reefers on board or the bow thruster. The estimated fuel efficiency has then been divided by the amount of containers on board to create an indicator for the energy efficiency of the vessel. For the calculation a uniform slot utilization of 85% is assumed. The calculated energy efficiency is determined by a mass of fuel in kg that is required to transport 1 TEU by a distance of 1 nm. This indicator is denominated as kg/TEUmile.

The result of this estimation is depicted in Fig. 5. The diagram shows that clear improvements with regard to energy efficiency of ULCVs were made over time. In Sect. 2.2 it was mentioned that with the commissioning of the Maersk E-class in 2006 a new chapter in ship size growth was opened. At the time of its commissioning, the energy efficiency of the E-class was significantly improved compared to the largest vessel that was commissioned 10 years earlier. Because the estimation does not include other factors such as the operation of auxiliary engines, boilers or the amount of reefer containers, the energy efficiency only depends on the main engine and the operating speed as well as the slot utilization and reflects the energy efficiency of the ship at sea. Because a uniform slot utilization has been assumed, an increase or reduction would lead to a more or less uniform increase or reduction of the energy efficiency among the vessels.

Figure 6 shows the improvements of energy efficiency in retrospective. To even out fluctuations, the energy efficiency of the analysis has been grouped into tranches of 5 years and their average energy efficiency was determined. Then the fuel consumption of the latest tranche, which contains the ULCV from 2015 to 2019, was compared to earlier tranches of ships. The figure shows that a ULCV from the latest tranche uses about 45% less fuel per TEUmile than a vessel from the first tranche, in which the Maersk C-class, which is not considered as a ULCV, was

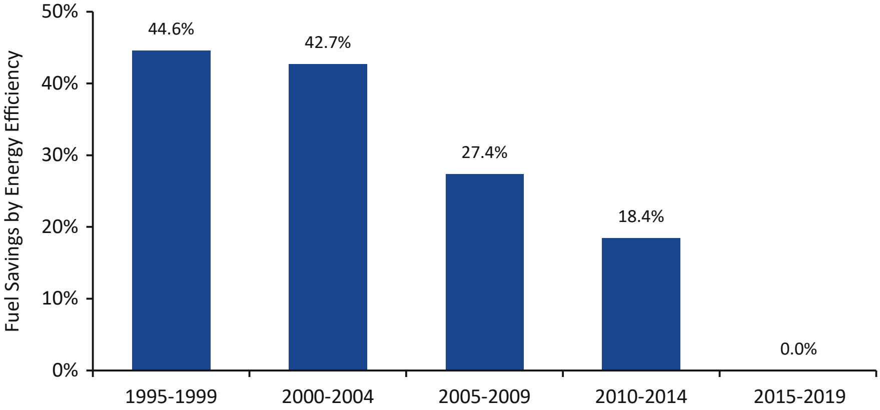


Fig. 6 Energy efficiency of a modern day ULCV compared to older vessels. Energy efficiency of different tranches is shown. Own diagram based on CRSL (2021a)

commissioned. Compared to the third tranche, which includes the Maersk E-class that coined the term 'ULCV', the last tranche still utilizes about 27% less fuel per amount of transport work.

Generally, there have been improvements to the energy efficiency of ULCVs that can be verified by an estimation of the energy efficiency. However, Fig. 5 also shows that improvements of the energy efficiency in the recent years have been achieved in rather small steps. Part of the reason is that the main engines are already highly optimized and achieve very low specific fuel oil consumptions, which is the mass of fuel that is required to produce a specific amount of energy. Due to this already high optimization of marine engines, the room for further improvements is small and operational measures such as weather routing or slow steaming become more relevant, because they still provide a leverage to the fuel consumption of a vessel (Faber et al. 2020). To further improve the fuel efficiency, operational measures or technical solutions such as air lubrication or wind assistance could be utilized. In the most recent ULCVs from 2020 and 2021 ship owners often implement dual fuel engines, that do not only use oils as a fuel but also Liquefied Natural Gas (LNG). While the operation of vessels with LNG results in lower CO_2-emissions, which gives the vessel a good EEDI rating, LNG is mostly composed of methane, a strong GHG, which can result in a harsh increase of CO_2e-emissions if leaked into the environment. The amount of this so-called methane slip that is the result of an incomplete combustion heavily depends on the engine that is installed. In the worst case, the use of LNG in dual fuel engines does result in significantly lower CO_2-emissions but in an increase of CO_2e-emissions. If the methane slip over the life cycle of LNG can be reduced or prevented, this fuel could provide a new leverage to further improve the efficiency of ULCVs (Pavlenko et al. 2020). While some shipping companies build their ships with LNG dual fuel engines (CMA CGM n.d.), others are using methanol (A.P. Moller - Maersk 2021) or even consider ammonia as potential future fuels (Lloyd's Register 2019), that could even be produced from renewable resources. In a recent study, the DNV GL considers several future fuels. Among LNG, methanol and ammonia from renewable sources, hydrogen, that is used in fuel cells, is part of the maritime future energy mix in some of the considered scenarios (DNV GL 2020).

4.3 Dredging for the Future

Dredging in ports is primarily done to accommodate large vessels and to allow large vessels to pass each other with a safe distance. ULCVs are one of the main reasons for dredging that is performed in ports all around the world. The environmental impacts that dredging has on the marine environment cannot be attributed solely to ULCVs but rather to the cascading effects with larger vessels replacing smaller units, as has been demonstrated in Sect. 2.2 (European Dredging Association n.d.).

In the context of ULCVs mostly two of these types of dredging are important. As has been discussed in Sect. 2.1, ships have always been getting bigger with ULCVs always being the peak of this evolution. While ports have often developed historically according to the needs of the time which maybe the near future in hindsight, in many cases they are not prepared for vessels the size of a modern day ULCV. To accommodate these large vessels the infra- and superstructure of the ports need to be modified as already discussed in Sect. 3.1. Modifications to the infrastructure, such as the enlargement or deepening of channels, berths or basins require dredging with the aim of creating or modifying infrastructure. This is called ‘capital dredging’. Due to different reasons, such as waves and tides, sediments are transported with the water and accumulate in channels, berths or basins of a port, where they can reduce the depth of a channel. To ensure that the port and the adjoining waters remain navigable, the so-called maintenance dredging is carried out at intervals, and the accumulated sediments are dredged (Ports Australia 2016; European Dredging Association n.d.).

Dredging is often considered as an interference into existing ecosystems. While it is obvious that dredging can be an issue to these ecosystems, the extent of the impact of a dredging measure is difficult to evaluate and the public discussion about it can lead to serious delays, especially when it evolves around capital dredging. A prime example for such a delay is the lengthy discussion about the dredging of the river Elbe in Hamburg (Germany), which started in 2002 and was ended by the approval of the works in 2020 (Grabbe 2017; Groll 2020; Merk et al. 2015). The effects of dredging on the marine environment are investigated intensively by many researchers. Generally speaking, the environment is affected in many different ways by dredging. During the dredging work itself, noise levels increase significantly above normal levels, which could be harmful especially to marine mammals (Todd et al. 2015). The process of dredging causes the so-called turbidity plumes, an increased amount of sediments in the water that have short-term effect on the quality of water (Ports Australia 2016). The turbidity plumes can negatively impact the light penetration of the water, which has the potential to degrade ecologically important habitats close to the dredging operation that are sensitive to any change of their environment. The extent of the effect on flora and fauna depends on the specific sensitivity of the specie, with some being able to adapt to changes and others not (Erftemeijer et al. 2012; Ports Australia 2016; Erftemeijer and Lewis 2006). Sediments can accumulate different toxins and pollutants over time. Dredging operations disturb these sediments and can release these toxins and pollutants into the water and surrounding habitats (Johnston 1981; Todd et al. 2015). Another impact is the change of the sediment structure due to dredging, which could become unsuitable as a habitat for some species. Again, the extent of the effect is varying on the predominant species of flora and fauna as well as the composition of the sediment before the dredging. Dredging can also have a positive short-term effect. Whirling up sediments can release nutrients into the water and increase the supply of food available for different species in an area, leading to a temporary rise of species abundance in the dredging area (Todd et al. 2015).

To ensure that the interference of marine dredging works on the environment is kept at a minimum, an environmental risk assessment is a crucial part of the planning stage. It shall be ensured that any unnecessary environmental impact is avoided and that any remaining impacts are mitigated and offset. Any impact that dredging has on water and sediment quality has to be evaluated. Additional aspects such as impacts or potential losses of marine flora and fauna, cultural heritage, fishing and shipping activities and any changes to current and water flows are to be considered (Ports Australia 2016). Concluding, studies do show that dredging negatively impacts marine life and the environment. It is indicated that the extent of these effects depends on the specific species, with some being more sensitive to disturbances than others (Erftemeijer and Lewis 2006; Erftemeijer et al. 2012; Todd et al. 2015; Johnston 1981). It is hinted that dredging can even have short-term positive impacts (Todd et al. 2015). In any case, negative effects outweigh the positive effects and it can take up to 10 years after the dredging operation until the area has fully recovered from the interference (Newell et al. 1998; Todd et al. 2015).

5 Conclusion

5.1 Various Effects of Ever Larger Ships

The discussion on seaport systems as well as on environmental considerations has shown that cost benefits realized at sea may come with negative effects in other external systems. Potential negative impacts on aquatic ecosystems and increasing logistical, infrastructural and superstructural challenges in the entire transport chain become more and more complex and extreme as the size and capacity of ULCVs further increase. Economic advantages for liner shipping companies through economies of scale therefore come at the expense of diseconomies of scale, which must be shouldered, to a large extent, by the general public.

ULCVs, meaning container ships that can no longer pass through the new locks of the Panama Canal due to their dimensions, have assumed a dominant role within the global container ship fleet since their introduction in 2006. While these vessels initially possessed nominal capacities in the range of 15,000 to 16,000 TEU, no ULCV with less than 19,000 TEU has been put into service since 2019. The current order book for new containerships indicates that only ULCVs with more than 23,000 TEU will become operational in the coming years and that a large part of the future growth in the total capacity of the global containership fleet will fall into this segment.

With regard to seaport systems, it becomes obvious that the rise of ULCVs has created a very demanding environment for port and terminal operators. While those vessels are often mentioned in connection with significant cost advantages achieved through 'economies of scale', it is argued that ports and terminals have rather become 'sweating ports' that suffer from 'diseconomies of scale' under the influence of large vessels (Haralambides 2019; Malchow 2017). On the one hand, this refers to necessary infra- and superstructure adaptions that are connected with huge investments, often subsidized by taxpayers' money. On the other hand, shipping companies as the terminals' main customers set increased operational performance requirements that can hardly be fulfilled at the moment and call for more and faster equipment as well as more flexibility. In connection with lower call frequencies and workload peaks as a result of increased call sizes, this environment is likely to lower the profitability of terminal operations. However, it is not guaranteed that the efforts to meet ULCV requirements will ultimately pay off by means of increased throughput, as shipping companies are in the process of reshaping their service structures that involve fewer ports of call within hub-and-spoke-networks.

It has become apparent that container ships come with various impacts on the environment. While these impacts are not unique to ULCV, it was demonstrated that these vessels provide an ecologic leverage over smaller units since they enable an increasing nominal capacity, while maintaining comparable outer dimensions. The ecological impact that a ULCV has can therefore be distributed among a higher number of containers. On a per-box-base the ecological impact is decreasing with the deployment of ULCV. This effect can best be described as 'ecologies of scale'. Hence, it could be demonstrated that the energy efficiency of ULCV has improved drastically over the last decades: due to improved engine efficiency and improved operational measures, a modern day ULCV is about 40% more energy efficient. Since engines have been highly optimized, further potential for energy efficiency gains from engine optimization is limited. Additional emission savings could be contributed from operational measures and additional technical solutions such as wind assistance or from future fuels. Even though it was shown that ULCVs do offer ecologic benefits mainly from improved machinery and 'ecologies of scale', dredging in port areas is mainly done to keep ports up with the increasing ship sizes. While it is hinted that capital dredging indeed can have short-term benefits on the marine environment, they are far outweighed by negative effects that often do have a long-term effect.

5.2 Economies and Diseconomies of Scale: The Discussion Goes into the Next Round

Due to the strong increase in orders, especially for ULCV new buildings since the fourth quarter of 2020, the future capacity growth of the world container fleet will be mainly driven by this largest ship segment in the coming years. It has been shown that the growth in size can in principle be continued and that the ULCV-class of MGX-24 ships could be surpassed by ships with up to 30,000 TEU, which could be significantly longer and wider than the largest ships in service today.

While consulting firms and economists have been grappling for some time with the question of how much and under what conditions further growth in the size of container ships can still lead to measurable economies of scale for liner shipping companies, diseconomies of scale of ULCVs have so far been described mainly in qualitative terms. Whereas negative ecological impacts, e.g. due to dredging, are difficult to quantify monetarily due to their long-term nature, this does not apply to ULCV-related modernization and maintenance costs of seaport network structures. Especially in this area, further research is needed to better estimate monetary consequences of ULCVs and possibly even larger units in the future. If, when and to which magnitude gigantism in container shipping will be carried on depends largely on whether the liner shipping companies will still be able to profit from economies of scale with ever larger ships and to what extent external diseconomies of scale are tolerated by the respective actors, such as the general public. Today, some voices even call for a limitation of vessel sizes in order to prevent ports from becoming unprofitable and public bodies from excessive spending in infrastructure that comes without significant compensation by vessel operators (WorldCargo News 2021b; Malchow 2017).

References

A.P. Moller - Maersk.: A.P. Moller - Maersk joins Methanol Institute. https://www.maersk.com/news/articles/2021/03/24/ap-moller-maersk-joins-methanol-institute (2021). Accessed 21 June 2021

Alphaliner: The next step for jumbo boxships: the ‘Megamax-24’. Alphaliner Weekly Newsletter. **2017**(45), 19–22 (2017)

Alphaliner: The ‘gigamax’: would a ‘post-megamax’ ship make any sense? Alphaliner Weekly Newsletter. **2021**(22), 1–5 (2021)

American Bureau of Shipping.: ABS Record Eugen Maersk. https://www.eagle.org/portal/#/absrecord/details. (2021). Accessed 23 April 2021

Andrés, S., Piniella, F.: Aids to navigation systems on inland waterways as an element of competitiveness in ULCV traffic. J. Traffic. Transp. Eng. **7**, 1–18 (2017). https://doi.org/10.7708/ijtte.2017.7(1).01

Böse, J.W.: General considerations on terminal planning, innovations and challenges. In: Böse, J.W. (ed.) Handbook of Terminal Planning, pp. 3–27. Springer International Publishing, Cham (2020)

Carbone, V., Martino, M.D.: The changing role of ports in supply-chain management: an empirical analysis. Marit. Policy Manag. **30**, 305–320 (2003). https://doi.org/10.1080/0308883032000145618

Cariou, P., Cheaitou, A.: Cascading effects, network configurations and optimal transshipment volumes in liner shipping. Marit. Econ. Logist. **16**, 321–342 (2014). https://doi.org/10.1057/mel.2014.4

Čerin, P., Beškovnik, B. Impact of container vessels capacity 15.000 TEU and more on existing container terminals. J. Mar. Transp. Sci. **3**, 49–61 (2020). https://doi.org/18048/2020.00.03

CMA CGM.: The CMA CGM JACQUES SAADE, the world's first 23,000 TEU powered by LNG. https://cmacgm-group.com/en/launching-cmacgm-jacques-saade-world%27s-first-ultra-large-vessel-powered-by-lng (n.d.). Accessed 21 June 2021

CRSL.: Shipping Intelligence Network (SIN). Clarksons Research Services Limited. https://sin.clarksons.net/ (2021a). Accessed 24 June 2021

CRSL.: World Fleet Register (WFR). Clarksons Research Services Limited. https://www.clarksons.net/wfr/ (2021b). Accessed 24 June 2021

de Haas, J.: Gard Guidance on Freight Containers. Gard, Arendal (2016)

Det Norske Veritas AS: Container Ship Update. Det Norske Veritas AS, Høvik, Norway (2013)

DNV GL. Maritime Forecast to 2050 (2020)

Ducruet, C., Notteboom, T.: Developing liner service networks in container shipping. In: Song, D.-W., Panayides, P.M. (eds.) Maritime Logistics: A Complete Guide to Effective Shipping and Port Management, pp. 77–100. Kogan Page, London (2012)

Erftemeijer, P.L., Lewis, R.R.R.: Environmental impacts of dredging on seagrasses: a review. Mar. Pollut. Bull. **52**, 1553–1572 (2006). https://doi.org/10.1016/j.marpolbul.2006.09.006

Erftemeijer, P.L., Riegl, B., Hoeksema, B.W., Todd, P.A.: Environmental impacts of dredging and other sediment disturbances on corals: a review. Mar. Pollut. Bull. **64**, 1737–1765 (2012). https://doi.org/10.1016/j.marpolbul.2012.05.008

European Commission: Integrating Maritime Transport Emissions in the EU's Greenhouse Gas Reduction Policies: COM(2013) 479 Final. European Commission, Brussels, Belgium (2013)

European Commission: 2019 Annual Report on CO2 Emissions from Maritime Transport: SWD(2020) 82 Final. European Commission, Brussels, Belgium (2020)

European Dredging Association.: EuDA - About dredging - Definitions. https://european-dredging.eu/Definitions (n.d.). Accessed 8 June 2021

Faber, J., Hanayama, S., Zhang, S., Pereda, P., Comer, B., Hauerhof, E., et al.: Fourth IMO GHG Study 2020: Reduction of GHG Emissions from Ships, Delft (2020)

Ge, J., Zhu, M., Sha, M., Notteboom, T., Shi, W., Wang, X.: Towards 25,000 TEU vessels? A comparative economic analysis of ultra-large containership sizes under different market and operational conditions. Marit. Econ. Logist. (2019). https://doi.org/10.1057/s41278-019-00136-4

Grabbe, H.: Elbvertiefung: Alles im Fluss. https://www.zeit.de/2017/07/elbvertiefung-hamburg-hafen-pro-contra (2017). Accessed 14 June 2021

Groll, T.: Elbvertiefung: Bundesgericht genehmigt Ausbau der Elbe endgültig. https://www.zeit.de/politik/deutschland/2020-06/elbvertiefung-hamburg-bundesverwaltungsgericht-leipzig-umweltschutz-klage (2020). Accessed 14 June 2021

Hamburg Port Authority: Port Information Guide – Port of Hamburg. Hamburg Port Authority, Hamburg (2018)

Hansestadt Bremisches Hafenamt: Port Information Guide – Port of Bremerhaven. Hansestadt Bremisches Hafenamt, Bremerhaven (2021)

Haralambides, H.: Gigantism in container shipping, ports and global logistics: a time-lapse into the future. Marit. Econ. Logist. **21**, 1–60 (2019). https://doi.org/10.1057/s41278-018-00116-0

Heaney, S., Poore, D., Behera, S., Dixon, M., Kumar, A., Tomar, R., et al.: Container: Forecaster & Annual Review 2020/21. Quarter 3 September 2020, London (2020)

Hill, D.: ZPMC Delivers 2 STS Cranes to RWG in Rotterdam. Heavy Lift News.https://www.heavyliftnews.com/zpmc-delivers-2-sts-cranes-to-rwg-in-rotterdam/ (2020a, April 23). Accessed 12 June 2021

Hill, D.: Zhen Hua 26 unloads 2 ZPMC STS Cranes in Hamburg. Heavy Lift News. https://www.heavyliftnews.com/zhen-hua-26-unloads-2-zpmc-sts-cranes-in-hamburg/ (2020b, May 11). Accessed 12 June 2021

ILO (International Labour Organization): Maritime Labour Convention. ILO, Geneva (2006)

IMO.: IMO Action to Reduce GHG Emissions from International Shipping: Implementing the Initial IMO Strategy on Reduction of GHG Emissions from Ships (n.d.)

IMO (International Maritime Organization): International Convention on the Control of Harmful Anti-Fouling Systems on Ships, 2001. IMO, London (2001)

IMO (International Maritime Organization).: IMO Environment Committee approves amendments to cut ship emissions. https://www.imo.org/en/MediaCentre/PressBriefings/pages/42-MEPC-short-term-measure.aspx (2020a). Accessed 11 June 2021

IMO (International Maritime Organization): International Convention for the Safety of Life at Sea, 1974: SOLAS 1974 (2020th ed., IMO Publication). International Maritime Organization, London (2020b)

IMO (International Maritime Organization).: IMO working group agrees further measures to cut ship emissions. https://www.imo.org/en/MediaCentre/PressBriefings/pages/36-ISWG-GHG-7.aspx (2020c). Accessed 21 May 2021

IMO (International Maritime Organization).: Further shipping GHG emission reduction measures adopted. https://www.imo.org/en/MediaCentre/PressBriefings/pages/MEPC76.aspx (2021). Accessed 24 June 2021

JOC Group: Berth Productivity: The Trends, Outlook and Market Forces Impacting Ship Turnaround Times. JOC Group (2014)

Johnston, S.A.: Estuarine dredge and fill activities: a review of impacts. Environ. Manag. **5**, 427–440 (1981). https://doi.org/10.1007/BF01866820

Koutroukis, G., Papanikolaou, A., Nikolopoulos, L., Sames, P., Köpke, M.: Multi-objective optimization of container ship design. In: Soares, C., Peña, F. (eds.) Developments in Maritime Transportation and Exploitation of Sea Resources, pp. 477–489. CRC Press, Boca Raton (2013)

Lane, A., Moret, C.: The Impact of Ever Larger Vessels on Terminals. CTI Consultancy, Singapore (2014)

Lian, F., Jin, J., Yang, Z.: Optimal container ship size: a global cost minimization approach. Marit. Policy Manag. **46**, 802–817 (2019). https://doi.org/10.1080/03088839.2019.1630760

Liebherr Group.: Crane Sizes and Technical Specifications. https://www.liebherr.com/en/deu/products/maritime-cranes/port-equipment/container-bridges/ship-to-shore-container-cranes.html#!/mcm-sts-accordian-start=mcm-sts-accordion-crane-sizes-start (2021). Accessed 11 June 2021

Lloyd's Register.: Industry project to design ammonia-fuelled 23k ULCS concept. https://www.lr.org/en/latest-news/aip-ammonia-fuelled-ulcs/ (2019). Accessed 21 June 2021

Malchow, U.: Growth in containership sizes to be stopped? Marit. Bus. Rev. **2**, 199–210 (2017). https://doi.org/10.1108/MABR-01-2017-0001

Mallin, P., To, A.: Ports Must Adapt to Mega-Ships to Remain Competitive. Maritime Information Services, London (2016)

Martin, J., Martin, S., Pettit, S.: Container ship size and the implications on port call workload. Int. J. Shipp. Transp. Logist. **7**, 553 (2015). https://doi.org/10.1504/IJSTL.2015.072016

MEPC (Marine Environment Protection Committee).: International Convention for the Prevention of Pollution from Ships, MARPOL 73/78 (1973)

MEPC (Marine Environment Protection Committee). Measures to Control Potential Adverse Impacts Associated With use of Tributyl Tin Compounds in Anti-Fouling Paints: MEPC.46(30) (1990)

MEPC (Marine Environment Protection Committee).: Shipping noise and marine mammals: MEPC 57/INF.4 (2007)

MEPC (Marine Environment Protection Committee).: Guidelines for the Development of the Inventory of Hazardous Materials: MEPC.197(62) (2011a)

MEPC (Marine Environment Protection Committee).: Inclusion of regulations on energy efficiency for ships in MARPOL Annex VI: MEPC.203(62) (2011b)

MEPC (Marine Environment Protection Committee).: Guidelines for Safe and Environmentally Sound Ship Recycling: MEPC.210(63) (2012a)
MEPC (Marine Environment Protection Committee).: Guidelines for the Authorization of Ship Recycling Facilities: MEPC.211(63) (2012b)
MEPC (Marine Environment Protection Committee).: Guidelines for the Development of Garbage Management Plans: MEPC.220(63) (2012c)
MEPC (Marine Environment Protection Committee).: Guidelines for the Inspection of Ships Under the Hong Kong Convention: MEPC.223(64) (2012d)
MEPC (Marine Environment Protection Committee).: Guidelines for the Survey and Certification of Ships Under the Hong Kong Convention: MEPC.222(64) (2012e)
MEPC (Marine Environment Protection Committee).: Standard Specification for Shipboard Incinerators: MEPC.244(66) (2014a)
MEPC (Marine Environment Protection Committee).: Guidelines for the Reduction of Underwater Noise from Commercial Shipping to address adverse Impacts on Marine Life: MEPC.1/Circ.833 (2014b)
MEPC (Marine Environment Protection Committee).: Guidelines for the Development of the Inventory of Hazardous Material: MEPC.269(68) (2015)
MEPC (Marine Environment Protection Committee).: Effective Date of Implementation of the Fuel Oil Standard in Regulation 14.1.3 of MARPOL Annex VI: MEPC.280(70) (2016)
MEPC (Marine Environment Protection Committee).: Guidelines for the Implementation of MARPOL Annex V: MEPC.295(71) (2017)
MEPC (Marine Environment Protection Committee).: Initial IMO Strategy on Reduction of GHG Emissions From Ships: MEPC.304(72) (2018)
MEPC (Marine Environment Protection Committee).: Guidelines for Consistent Implementation of the 0.50% Sulphur Limit under MARPOL Annex VI: MEPC.320(74) (2019)
MEPC (Marine Environment Protection Committee).: Procedures for sampling and verification of the sulphur content of fuel oil and the Energy Efficiency Design Index (EEDI): MEPC.324(75) (2020)
Merk, O.: Container Ship Size and Port Relocation: Discussion Paper, Paris (2018)
Merk, O., Busquet, B., Aronietis, R.: The Impact of Mega-Ships: Case-Specific Policy Analysis (2015)
Mongelluzzo, B. LA-LB proves ability to handle mega-ships, but landside pressures remain. https://www.joc.com/port-news/terminal-operators/la-lb-proves-ability-handle-mega-ships-landside-pressures-remain_20200623.html (2020). Accessed 21 June 2021
MSC (Maritime Safety Committee).: Adoption of Amendments to the International Convention for the Safety of Life at sea, 1974, as Amended: MSC.338(91) (2012a)
MSC (Maritime Safety Committee).: Adoption of the Code on Noise Levels on Board Ships: MSC.337(91) (2012b)
Musso, E., Sciomachen, A.: Impact of megaships on the performance of port container terminals. Marit. Econ. Logist. **22**, 432–445 (2020). https://doi.org/10.1057/s41278-019-00120-y
Newell, R.C., Seiderer, L.J., Hitchcock, D.R.: The impact of dredging works in coastal waters: a review of the sensitivity to disturbance and subsequent recovery of biological resources on the sea bed. Oceanogr. Mar. Biol. Annu. Rev. **36**, 127–178 (1998)
Notteboom, T.: The adaptive capacity of container ports in an era of mega vessels: the case of upstream seaports Antwerp and Hamburg. J. Transp. Geogr. **54**, 295–309 (2016)
Ozbas, B., Spasociv, L., Campo, M., Besenski, D.: Analyses of interactions between the marine terminal and highway operations (2014)
Panama Canal Authority. OP notice to shipping no. N-1-2021: Vessel Requirements. Panama Canal Authority (2021)
Park, N.K., Suh, S.C.: Tendency toward mega containerships and the constraints of container terminals. J. Mar. Sci. Eng. **7**, 131 (2019). https://doi.org/10.3390/jmse7050131
Pavlenko, N., Comer, B., Zhou, Y., Rutherford, D.: The climate implications of usinrijg LNG as a marine fuel (2020)

PEMA (Port Equipment Manufacturers Association).: 7th Annual Ship-to-Shore Crane Survey: Global Deliveries 2017 (2018)

PIANC.: Harbour approach channels design guidelines: report no 121–2014 (2014)

Ports Australia: Environmental Code of Practice for Dredging and Dredged Material Management. Ports Australia, Sydney (2016)

Psaraftis, H.N., Kontovas, C.A.: Decarbonization of Maritime transport: is there light at the end of the tunnel? Sustainability. **13**, 237 (2021). https://doi.org/10.3390/su13010237

Rijsenbrij, J., Wieschemann, A.: A powerful union for sustainable container terminal design. In: Böse, J.W. (ed.) Handbook of Terminal Planning. Springer International Publishing, Cham (2020)

Rodrigue, J.-P. (2021). Channel Depth at Major North American Container Ports | The Geography of Transport Systems. https://transportgeography.org/contents/chapter6/port-terminals/channel-depth-ports-north-america/. Accessed 28 June 2021

Russell, D., Ruamsook, K., Roso, V.: Managing supply chain uncertainty by building flexibility in container port capacity: a logistics triad perspective and the COVID-19 case. Marit. Econ. Logist. (2020). https://doi.org/10.1057/s41278-020-00168-1

Rutherford, D., Mao, X., & Comer, B. (2020). Potential CO2 Reductions under the Energy Efficiency Existing Ship Index

Sahoo, P.K.: Principles of Marine Vessel Design: Concepts and Design Fundamentals of Sea Going Vessels. World Scientific Publishing Company, Singapore (2021)

Scheldecoördinatiecentrum Vlissingen.: VTS-Scheldt - Tijpoort in relatie tot het toelatingsbeleid. https://www.vts-scheldt.net/nautisch/tijpoorten (2021). Accessed 16 June 2021

Schlich, E., Fleissner, U.: The ecology of scale: assessment of regional energy turnover and comparison with global food (5 pp). Int. J. Life Cycle Assess. **10**, 219–223 (2005). https://doi.org/10.1065/lca2004.09.180.9

Schönknecht, A.: Maritime Containerlogistik: Leistungsvergleich von Containerschiffen in intermodalen Transportketten. Springer, Berlin/Heidelberg (2009)

Soderberg, E.: The effect of ultra large container vessels on cranes and infrastructure (2017)

Specht, P., Kotzab, H., Arendt, F. Uncertainty affecting seaport-hinterland coordination: the case of information on cargo availability. In: 19th International Conference on Transport Science, Portoroz, Slovenia pp. 313–319 (2020)

Stopford, M.: Maritime economics. Routledge, London, New York (2009)

Stoschek, O., Leschka, S., Hein, C., Brüning, A.: Optimizing pier structures using dynamic mooring forces modelling. PIANC-World Congress Panama City, Panama (2018)

Sys, C., Blauwens, G., Omey, E., van de Voorde, E., Witlox, F.: In search of the link between ship size and operations. Transp. Plan. Technol. **31**, 435–463 (2008). https://doi.org/10.1080/03081060802335109

THB.: TEU-Upgrade für ‚Emma Maersk“ und Co. https://www.thb.info/rubriken/detail/news/teu-upgrade-fuer-emma-maersk-und-co.html (2018). Accessed 23 April 2021

The Loadstar.: Measuring port performance (2015)

Todd, V.L.G., Todd, I.B., Gardiner, J.C., Morrin, E.C.N., MacPherson, N.A., DiMarzio, N.A., et al.: A review of impacts of marine dredging activities on marine mammals. ICES J. Mar. Sci. **72**, 328–340 (2015). https://doi.org/10.1093/icesjms/fsu187

Trivyza, N.L., Rentizelas, A., Theotokatos, G.: A comparative analysis of EEDI versus lifetime CO2 emissions. J. Mar. Sci. Eng. **8**, 61 (2020). https://doi.org/10.3390/jmse8010061

van Ham, J.C.: The feasibility of mega container vessels. European Transport\Trasporti Europei. **25**(25–26), 89–98 (2005)

Vantorre, M., Candries, M., Verwilligen, J.: Optimisation of tidal windows for deep-drafted vessels by means of a probabalistic approach policy for access channels with depth limitations (2014)

Wijnolst, N., Scholtens, M., Waals, F.: Malacca-Max: the Ultimate Container Carrier Delft Marine Technology Series, vol. 14. Delft Univ. Press, Delft (1999)

WorldCargo News.: Triple E ships - 6000 moves in 24 hours. https://www.worldcargonews.com/news/news/triple-e-ships%2D%2D-6000-moves-in-24-hours-29178 (2011). Accessed 21 June 2021

WorldCargo News.: Hapag-Lloyd to start calling JWP Wilhelmshaven? https://www.worldcargonews.com/news/news/hapag-lloyd-to-start-calling-jwp-wilhelmshaven-66373 (2021a). Accessed 21 June 2021

WorldCargo News.: Staying in the game as ships grow. https://www.worldcargonews.com/in-depth/staying-in-the-game-as-ships-grow (2021b). Accessed 24 June 2021

Modeling Individualized Sustainable Last Mile Logistics

Markus Trapp, Sandra Luttermann, Daniel Rippel,
Herbert Kotzab, and Michael Freitag

Abstract The online grocery trade has received an additional boost from the Covid pandemic. The delivery of such purchases places particular demands on last mile logistics since consumers demand more and more individualized delivery options, e.g., regarding the delivery arrival or the type of transport. At the same time, many consumers are becoming more environmentally conscious, so there is a need to examine further how this particular consumer behavior affects the sustainability of deliveries. This paper develops and presents a simulation model, which considers grocery delivery under different framework conditions. The examined scenarios show that a change in consumer behavior directly impacts last mile logistics systems, mainly by increasing the total number of orders and a slight reduction in emissions through improved vehicle utilization. Nevertheless, the results show that without sufficiently high utilization of delivery vehicles, shopping trips by private car may cause fewer emissions.

M. Trapp (✉) · D. Rippel
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany
e-mail: tap@biba.uni-bremen.de; rip@biba.uni-bremen.de

S. Luttermann · H. Kotzab
Faculty of Business Studies and Economics, University of Bremen, Bremen, Germany
e-mail: sandra.luttermann@uni-bremen.de; kotzab@uni-bremen.de

M. Freitag
BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Bremen, Germany

Faculty of Production Engineering, University of Bremen, Bremen, Germany
e-mail: fre@biba.uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_13

1 Introduction

Nowadays, more and more consumers are using e-commerce for their purchases, which leads to increased delivery volumes, especially in urban areas. Thereby, we observe that individual order frequencies trigger large numbers of individual shipments of single orders with single packages. Consequently, the utilization of transport means decreases (Pronello et al. 2017; Schnedlitz et al. 2013). From a consumer's point of view, there is also a rising desire for more individuality in the delivery of online ordered goods (Witten and Schmidt 2019). Consumers expect flexible deliveries as well as delivery services that are individually tailored to their lifestyles. This refers particularly to timely delivery of ordered goods, direct deliveries to the front door, and well-determined delivery dates (Wegner 2019). At the same time, it is apparent that the probability of personal receipt and thus successful delivery tends to decline.

Furthermore, consumers are becoming more environmentally aware and want to purchase in an environmentally sustainable manner (Ternès et al. 2015). However, this conflicts with increasing traffic volumes that are generated by e-commerce. Thereby, the so-called last mile often turns out to be inefficient, costly, and environmentally damaging (Gevaers et al. 2011). Especially urban delivery areas are challenged by more traffic congestions, air and noise pollution, traffic accidents, and greenhouse gas emissions. These new complex logistics structures put retailers as well as logistics service providers under great cost pressure (Savelsbergh and Van Woensel 2016). Accordingly, alternative concepts for the last mile need to be developed for the future, which are both, sustainable as well as individual.

This paper examines different delivery options for online grocery retailing that allow for individual adjustments of end users according to their preferences. Therefore, we want to answer the following research question:

How do consumer choices in regards to delivery options of online purchased groceries affect the sustainability of last mile logistics?

In order to answer the research question, different delivery scenarios are developed and evaluated with regard to their sustainable impact.

The remainder of the paper is as follows. After having presented the motivation of this research and the research question, Sect. 2 shows theoretical notions on online grocery retailing, last mile logistics and its delivery options, as well as last mile related modeling is presented. Section 3 includes an explanation of the methodology of simulation modeling anda description of the processes. This is followed by the presentation and discussion of the identified results in Sect. 4. The paper ends in Sect. 5 with a conclusion, limitations, and an outlook for further research.

2 Theoretical Background

2.1 Online Grocery Retailing and Its Logistics Challenges

With an annual growth of around 60% between 2019 and 2020, the increase of online grocery retailing was significantly higher than the entire retail sector in the online business (HDE Online-Monitor 2021). Despite this significant growth rate, the market share of online grocery retailing is still relatively low, as only two percent of all grocery items are purchased online (HDE Online-Monitor 2021).

This is due to the many obstacles that make online grocery retailing more difficult than stationary retailing. These include issues of missing quality checks of purchased products during the purchasing process or the challenges of delivering fresh or frozen goods that trigger concerns among consumers (Deges and Speckmann 2020). In addition, the overall competition is considered to be extremely high due to a very good local supply situation through stationary retailing, the habitual structures of consumers, low margins and a high price sensitivity of consumers, who are often not willing to pay additional delivery costs (Heinemann 2020).

The supply chains for groceries differ significantly when stationary retailing is compared with online retailing. In conventional supply chains of stationary retailing, goods are delivered from manufacturers to the central distribution centers of retailers, from where they are forwarded to the local stores. In online supply chains, these process steps between manufacturer and end consumers can be reduced (Ehrler et al. 2019). End users buy and receive their goods either directly from the producer, similar to traditional market stalls, or from wholesalers or retail distribution centers (Seidel et al. 2016). In stationary retailing, ultimate logistics operations such as picking, packing, and delivery of individual orders are handled by consumers themselves, whereas in online grocery retailing, these fulfillment operations are transferred back to retailers and/or manufacturers (Yumurtacı Hüseyinoğlu et al. 2020; Kämäräinen and Punakivi 2002).

Furthermore, many logistics processes in the grocery retailing sector are subject to special legal regulations and controls in the areas of food hygiene, temperature safety, data protection and product traceability (Dworak and Burdick 2002; Vahrenkamp et al. 2012; Arnold et al. 2008). For example, legal requirements stipulate that the end consumer must be guaranteed that groceries remain fresh until the best-before date (Hagenmeyer 2006). This is particularly relevant for fresh products such as meat and sausage products as well as fruit and vegetables (Vahrenkamp et al. 2012). Consequently, an uninterrupted cold chain between producers and consumers is essential.

2.2 Modeling Last Mile Logistics and Its Delivery Options

2.2.1 Last Mile Definition and Delivery Options

The last mile is this part of a supply chain in which an order is delivered from the last distribution center, collection point or local warehouse to its destination (Umundum 2020). The final destination can be a private end user or a company (Clausen et al. 2016). Various players, such as manufacturers, retailers, and logistics service providers, are involved in managing the last mile. The player with the highest bundling potential can manage the last mile most efficiently and cost-effectively; today, this is most likely to be logistics service providers (Brabänder 2020).

The last mile offers a variety of delivery options, which differ significantly in terms of sustainability impact and at the same time offer the possibility of considering consumer-specific delivery choices. This includes the question of which means of transport is used to cover the last mile. This can be done, for example, by conventional vehicles with combustion engines or electrically powered vehicles such as e-cars or e-bikes. Cargo bikes are also becoming increasingly important, especially in hard-to-reach areas in city centers (Saenz et al. 2016). Furthermore, there are many design options in terms of speed of delivery. Consumers demand short delivery times in particular so that same-day or next-day delivery is becoming increasingly popular in online retailing (Winkenbach and Janjevic 2018). These requirements pose major problems for logistics service providers, as short delivery times lead to rising costs, make route planning more complicated, reduce vehicle utilization, and increase CO_2 emissions per package (Witten and Schmidt 2019). Also, different time slots for delivery can be offered to match individual consumer preferences. The selection of a time of the day is intended to ensure that the consumer is present at the time of delivery (Agatz et al. 2008). So far, the choice for a specific delivery time is limited, and the times for delivery specified by logistic service providers often collide with consumers' work schedules (Grant et al. 2014). Since successful first-time delivery has a strong influence on a sustainable delivery, there is great potential for logistics service providers to adapt their delivery times to the preferences of consumers.

Finally, different delivery locations can be adjusted to consumer preferences and have significantly different effects on CO_2 emissions during delivery. Consumers clearly prefer home delivery, which is the most common delivery option. The problem here, however, is that consumers often are often not present at the time of delivery, which triggers further delivery attempts and additional trips. There are concepts to avoid this problem and enable more sustainable delivery, such as click-and-collect, which involve self-collection by the consumer from the store or a packing station (Mangiaracina et al. 2019; Schnedlitz et al. 2013).

2.2.2 Modeling the Last Mile

Modeling last mile logistics often focuses on finding better algorithms for scheduling and clocking delivery trips for a wide variety of scenarios. It is considered as a special sort of a Vehicle Routing Problem (VRP), as last mile deliveries are usually constrained by capacity limits imposed by the vehicles and given time windows. Consequently, it is often referred to as a (Capacitated) Vehicle Routing Problem with Time Windows (VRPTW). Thereby, the focus is often on the reduction of vehicles. However, Ombuki et al. (2006) developed multi-objective genetic algorithms, where the number of vehicles and the incurred costs is equally important.

Depending on the application, almost any number of constraints can be added to the basic VRP model. For example, Shen et al. (2018) included the evolution of CO_2 emissions costs for a system with multiple depots and time windows (MDOVRPTW). They showed that adapting the delivery strategy to emissions regulations (costs and quotas) decreases overall costs. Current research also investigates the use of electric vehicles, where Caggiani et al. (2021) extend the VRPTW by including aspects of staggered delivery and the need to (partially) charge vehicles (2E-EVRPTW-PR). They conclude that despite these additional constraints, zero-emission strategies for the last mile can be worthwhile.

Also consumer accessibility impacts last mile transportation costs, which Özarık et al. (2021) have investigated by incorporating information on whether consumers are accessible into route planning and show that this can reduce overall costs.

Another important parameter that needs to be incorporated in a modeling approach refers to the problem that the adaptation of logistics processes or conversion of a vehicle fleet to zero-emission drives generates costs that cannot easily be passed on to consumers. Here, Hagen and Scheel-Kopeinig (2021) investigated the acceptance and willingness-to-pay (WTP) of consumers based on a central last mile micro depot (CMD). Their results show that many respondents are interested in using such a last mile service, but only a small proportion is willing to pay more for this additional service with potentially lower emissions.

3 Developing a Simulation Model for Individualized Sustainable Last Mile Logistics

3.1 Modeling Individualized and Sustainable Last Mile Logistics

Basically, individualization is used to achieve higher consumer loyalty by increasing the individual benefits for the end user (Reichwald et al. 2009). In the manufacturing industry, modularization, and postponement strategies already offer consumers many opportunities for individualization (Kölmel et al. 2019; Shaik et al. 2015). In contrast, the individualization of logistics in the last mile is hardly pronounced.

In accordance with Gausmann (2009), we understand consumer individuality as the end consumer's ability to make an individual choice from several logistics options. A choice is given if at least two different alternatives are offered to the consumer. With regard to the delivery options of groceries ordered online, such individualization can consist of the end user making a choice from various delivery dates, delivery locations, means of transport or transport packaging.

In our understanding, sustainability in last mile logistics refers primarily to the ecological dimension. Ecological sustainability is given when natural raw materials are only used under consideration of their natural recovery (Reichert et al. 2018). Against this background, we consider all raw materials and the emissions associated with their consumptions that are used and generated during the entire last mile transport process. This includes CO_2 emissions during loading, transport, and delivery of groceries ordered online. For example, we are investigating the extent to which CO_2 emissions change when different means of transport are used for delivery.

Based on this, we developed a simulation model for the Bremen area to investigate the impact of different logistics options for the delivery of online ordered groceries and their interaction on the respective environmental impact. In contrast to other approaches, the focus is not improving the routing and scheduling algorithm but an investigation of how varying specific logistics options change the outcomes.

Therefore, the model investigates the impact of individual consumers' selected logistics options in terms of dependent KPIs, such as individual CO_2 emissions. A model with public average values for consumers' shopping behavior from Germany was parameterized as a reference marker in the first step. We then created a second scenario using actual consumer preferences and shopping behavior from a survey. By comparing these two scenarios, we investigate whether and to what extent individual consumer preferences in the delivery of groceries ordered online have an impact on CO_2 emissions and other KPIs.

3.2 General Process Description

The simulation represents the following process: Consumers place orders in a continuous process according to their characteristics, which may differ due to the logistics options chosen. The orders received throughout the day are collected. Every time one or more vehicles becomes idle, the simulation tries to schedule all of today's remaining orders and to add as many optional orders (e.g., orders that could be delivered over the next five days). The applied routing tries to ensure that both the consumers' wishes can be fulfilled while obtaining a high vehicle utilization. In this first step, two vehicles with different engine types are available; in the further course, fleets with only one engine type are considered. In all cases, each route starts and ends at the main depot. Figure 1 illustrates this process in a schematic way.

The simulation collects KPIs, such as the total weight, the number of consumers or the distance driven per tour for each tour and individual order. The simulation

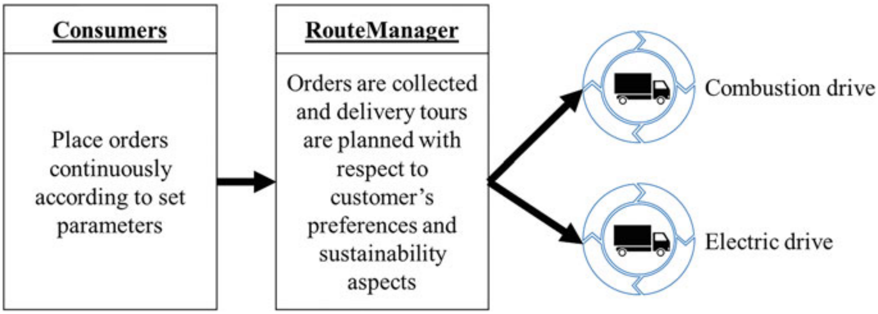


Fig. 1 Process model from ordering to delivery

calculates the individual environmental impact of consumer-related emissions using these KPIs. A detailed description of this approach is given in the following sections.

3.3 Simulation Model Description

Our simulation model was created with the AnyLogic software that uses Java as the programming language and provides an internal OpenStreetMaps GIS Map for routing and visualization. As the routing problem, a Capacitated VRP with Time Windows (CVRPTW) was identified, the GraphHopper's jsprit Libraries and a pre-calculated distance matrix were used to solve it. For each scenario in every variation, five runs over a simulation time of 30 days were done.

The core elements of the simulation model include agents for consumers and vehicles, in which the ordering and delivery behavior are described, and agents who manage incoming orders and optimize the routes. All these elements are presented below.

3.3.1 Consumers and Transport Matrix

The simulation model uses a consumer matrix consisting of 200 consumers randomly distributed in Bremen city. The corresponding distance matrix was calculated a priori using street distances within a GIS map. The resulting transport network uses the OpenStreetMap function in AnyLogic to drive along real roads and not just weighted edges in an artificial network.

For our simulation, we include the following delivery options and their attributes:

Each consumer is assigned to one of the possible values for each attribute. The distribution among the different values is based on consumer-specific preferences from our primary and secondary data sources. These logistics options determine different consumer-specific values that are essential for the simulation. For example,

– Number of persons	(1, 2, 3, 4, 5)
– Order frequency	(once a week, twice a week)
– Deposit	(yes, no)
– Desired time slots	(07–14 h, 14–20 h, 07–20 h)
– Desired transport vehicle	(combustion, electric)
– Desired period	(same day, next day, within next five days)

the number of people, the order frequency, and the desire for storage (impersonal delivery) determine each order's weight and the respective time required for packing and unloading. Traffic volumes differ throughout the day. Therefore, if a desired time slot with higher traffic volume is chosen, deliveries take longer and fuel consumption increases due to more stop-and-go sequences.

3.3.2 Vehicles

Our system consists of two delivery vehicles. Depending on the scenario, there are two combustion vehicles, two electric or one each. These two types mainly differ in consumption-related emissions and reactions to different traffic events. The specific total consumption $\text{con}_{\text{total}}$, given in Eq. (1) is determined by the type of engine and its related base consumption con_{base} as well as by a load weight factor $\text{con}_{\text{weight}}$ and a traffic influencing factor $\text{con}_{\text{traffic}}$ that reflects, e.g., different speeds and frequent stops depending on the time of day.

$$\text{con}_{\text{total}} = \text{con}_{\text{base}} \cdot \left(\text{con}_{\text{weight}} + \text{con}_{\text{traffic}}\right) \tag{1}$$

Equation (1) considers that the proportion per individual load is higher with a small total load and shrinks with the increasing total load. Thus, the respective consumption (diesel or electricity) is calculated for each order and can be offset against the respective specific emission values. Combustion vehicles tend to have higher fuel consumption according to higher amounts of stop-and-go-frequencies, resulting in a lower average speed (Olivera et al. 2015).

3.3.3 Routing

For the simulation, a dynamic planning problem has to be solved since new orders can also arrive during a day, which in the most extreme case are intended to be executed for the current day and should accordingly also be scheduled for the day if possible. Therefore, the agent first generates a route for orders that need to be completed on the same day. Afterward, the agent incrementally adds orders that might be scheduled for today but also could be delivered the following days. This process ensures a high vehicle utilization and tries to keep the number of orders on consecutive days low. The agent selects these additional orders depending on their urgency, i.e., it first tries to add orders that would be due the next day, then the day after and so on. The agent terminates the planning when the resulting route would not cover all orders anymore and assigns the complete plan to the idle vehicles.

3.3.4 Calculating Consumer-Related Emissions

The distance, the respective speed, the general traffic situation, the transport weight and the vehicle's capacity utilization result in the corresponding consumption (in this case, diesel or electricity) and therefore determine the total CO_2 emissions. By multiplication with the respective specific emission factor, a total value with the unit kg CO_2eq. per tour can be determined.

To get a reasonable emission share for each order, neither the consumer-related transport weight share nor the driven distance are suitable as sole calculation factors. Since a fully-loaded vehicle has a higher consumption than an empty one, the consumer's position within the planned tour influences the resulting emissions. To take into account that consumers have no active influence on their position, an allocation method is chosen that puts the product of distance and delivery weight per order in relation to the sum of all orders per tour shown in Eq. (2). Therefore, values per consumer and the total sum are rather determined by the consumers' choice of logistic options than by the tour-planning algorithm setting the sequence.

$$\text{prop}_{\text{order}} = \frac{\text{distance}_{\text{order}} \cdot \text{delivery weight}_{\text{order}}}{\text{distance}_{\text{tour}} \cdot \text{delivery weight}_{\text{tour}}} \tag{2}$$

Using this method, orders with very little weight but long distances and orders with significant weight but short distances get their fair share of the total. Having calculated the individual proportion of the respected tour, the amount of emissions for each consumer in the unit kg CO_2eq. can be determined by applying the formula shown in Eq. (3):

$$\text{emission}_{\text{order}} = \text{prop}_{\text{order}} \cdot \text{emission}_{\text{totalTour}} \tag{3}$$

4 Results and Discussion

4.1 Experimental Design

As mentioned before, the focus lies on creating data on consumer-related choices of logistics options. Therefore, in the different scenarios, varying choice sets are adjusted, and the resulting data is analyzed.

Table 1 shows the consumer-related features, their corresponding option lists and the assigned distributions. For the baseline scenario, literature values were used for the two features *number of person groceries are ordered for* and *order frequency*

Table 1 Features and distributions for the baseline and the empirical numbers scenario

Feature	Option list	Assigned distributions in %	
		Baseline scenario	Scenario with empirical numbers
Number of person groceries are ordered for	(1, 2, 3, 4, 5)	(50, 31, 9, 7, 3)[a]	(35, 40, 13, 6, 6)
Order frequency per week	(once, twice, daily)	(69, 27, 4)[b]	(42, 54, 4)
Deposit	(yes, no)	(20, 80)	(20, 80)
Desired time slot	(07–22 h, 08–14 h, 14–20 h)	(20, 20, 60)	(20, 20, 60)
Desired transport	(Diesel, electric)	(50, 50)	(50, 50)
Desired period	(SameDay, NextDay, within 5 days)	(20, 50, 30)	(20, 50, 30)

[a] In accordance with State Statistical Office Bremen (2016)
[b] In accordance with Kläver et al. (2020)

per week according to the sources given. A total of 260 participants took part in our survey on grocery shopping behavior, from which the distributions shown for the two features are derived. The remaining distributions are based on assumptions made by the authors.

By comparing the results of our survey data with the data used for the baseline scenario, three interesting aspects appear:

1. There is a shift in how many people are assigned to one grocery purchase.
2. The number of shopping trips per week differs significantly.
3. The vast majority of grocery shopping trips is done by car.

Whereas the data from the State Statistical Office Bremen say that almost 50% of all orders are for one person, only 35% of the survey participants chose this answer. The most noticeable part of this shift can be found in the shopping for two people, from 31% in the baseline scenario to 40% in the empirical scenario. Only small changes can be seen in the numbers for more than two people.

The second significant difference relates to the number of shopping trips per week. Survey data from Kläver et al. (2020) state that almost 70% of all consumers do grocery shopping once, or even less, a week. In our survey, this value drops to only 42%, but at the same time, the number for doing grocery shopping twice a week doubles to 54%. In both cases, daily shopping is only done by 4% of the participants. Lastly, our survey shows that grocery purchases are made almost exclusively in stationary stores and that about 70% of the survey participants use their car for this purpose. Here, the question arises whether CO_2 emissions can be reduced if online orders are placed and delivered by a small number of vehicles instead of many consumers do shopping trips by car.

Using the two different data sets, conclusions can be drawn about how the changing consumer preferences affect the delivery system. However, at the same time, in order to investigate developments when different vehicle fleets are available, simulations of these fleets were changed in a third step. While both vehicle types are selected at 50% in the initial scenarios, 100% use of one vehicle type was simulated for comparison purposes.

4.2 Experimental Results

In Table 2, the average results of five simulation run for each scenario are presented, beginning with the mixed vehicle fleet, followed by the combustion drive only and the electric drive only scenario. Finally, in Fig. 2, a comparison of all scenarios based on the emissions per order is shown.

The results show that consumer preferences have a direct impact on the logistics system. However, as illustrated in Fig. 2, within each vehicle distribution scenario, the average emissions per order remain similar at around 4.3, 3, and 1.3 kg CO_2eq. despite the change in consumer behavior, that leads to an apparent increase of around 15% of orders for all vehicle distributions. Figure 2 also shows that the average results on the emissions per order are always at the same level, but the variances are clearly different. Since the combustion only fleets show greater

Table 2 Average results of five simulation runs for a mixed vehicle fleet, combustion vehicles and electric vehicles only

KPI	Baseline scenario	Scenario with empirical numbers
Average results of five simulation runs for a mixed vehicle fleet		
Emissions per order/kg CO_2eq.	3.09	3.05
Orders delivered	707.00	833.60
Tours done	99.60	99.00
Orders per tour	7.18	8.45
Delivery weight per tour/kg	101.56	157.62
Distance per order/km	14.31	14.34
Average results of five simulation runs for combustion vehicles only		
Emissions per order/kg CO_2eq.	4.36	4.24
Orders delivered	762.60	856.20
Tours done	105.00	99.20
Orders per tour	7.35	8.69
Delivery weight per tour/kg	119.95	147.97
Distance per order/km	13.42	13.06
Average results of five simulation runs for electric vehicles only		
Emissions per order/kg CO_2eq.	1.35	1.30
Orders delivered	705.20	833.20
Tours done	105.20	101.20
Orders per tour	6.76	8.29
Delivery weight per tour/kg	91.14	140.54
Distance per order/km	13.63	13.04

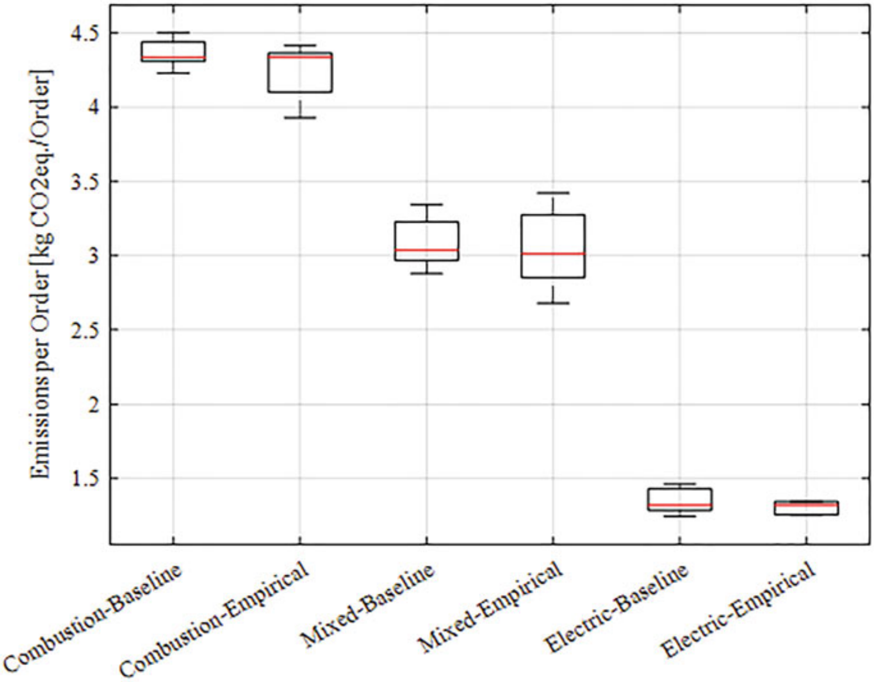


Fig. 2 Visualization of emissions per order for different vehicle distribution

fluctuations and higher values, these vehicles mainly shape the results for the mixed fleets.

Although more orders had to be delivered, the total amount of tours needed shrinks at least by a small factor. This development is also accompanied by an increase in delivery weight per tour. For both the mixed and the electric-only fleet, rises of more than 50% can be seen. However, for all vehicle distributions, it becomes obvious that there is a low utilization. Based on a maximum load capacity of around 1000 kg, the utilization in all baseline scenarios is around 10%. In comparison, in the empirical scenarios, it is still around almost 15%. Contrary to the weight per tour, for distances per tour, only small changes can be seen.

Finally, for all vehicle distributions, more efficient tours reduce emissions per order, which is achieved through the change in consumer behavior. Considering the assumptions made for this simulation model, full electrical fleets cause the smallest amount of emissions, but even a mixed fleet can significantly reduce environmental impacts.

4.3 Discussion on Results

The KPI *weight per trip* shows how consumer behavior impacts the environmental compatibility of the transport of online ordered groceries. Having the possibility to choose between several options, individualized deliveries are generated for consumers, but the providers of last mile logistics depend on a critical amount of consumers with similar selections made to generate environmentally friendly delivery tours. Although the planning algorithm tries to bundle as many orders as possible, our approach clearly shows the limits for optimizing a consolidation of the orders made by 125 active consumers with two vehicles only.

Since a majority of the respondents in our survey stated that purchases are made by car, it is worth estimating the emissions if all orders were picked up from stationary retail outlets by car. According to Schewelsky et al. (2020), CO_2 emissions per shopping trip in Germany amount to around 1.28 kg CO_2eq. In our simulation, comparable values result when delivery is made in electric vehicles only (1.35 kg CO_2eq. for the baseline scenario and 1.3 kg CO_2eq. for the scenario with empirical numbers). This highlights the potential positive sustainability impacts of using electric-powered vehicles to deliver groceries ordered online. In addition, traffic would be relieved if only two delivery vehicles were in use compared to everyone doing their grocery shopping by car. In contrast, the emissions in our simulation are more than twice as high when delivered with combustion vehicles. In this case, a shopping trip by car is less harmful to the environment than supplying the urban area from a single location that is positioned for practical reasons in the vicinity of the grocery store. In addition to the already mentioned low utilization of the delivery vehicles, a look at the distances per order also shows why the emission values are comparatively high. While in the simulation scenarios, distances of around 14 km per order have to be driven, the figure for private car trips is only around 5 km (Schewelsky et al. 2020). However, the results of our simulation do not allow a generally valid statement about how the emissions from shipping groceries ordered online behave compared to shopping trips by private car. Instead, it shows the need to investigate under which circumstances a breakeven is achieved, taking into account consumer behavior, fleet size, drive types, and possibly the number of locations concerning delivery areas and traffic situations and their interaction.

The fact that vehicle utilization has a significant impact on the efficiency and the outcomes of delivery routes is not new. However, this first simulation study shows that the delivery options directly influence the whole last mile logistics system. Therefore, to achieve a compromise between fulfilling consumers' wishes and planning and scheduling delivery routes with an environmental impact as little as possible, there is a need for a dynamic planning process. According to Özarık et al. (2021), who included data about the presence of consumers in their planning and scheduling system, information about all orders with specific restrictions like a particular time window or a delivery area could be considered. In that way, with every new order coming in, there would be a continuous calculation of transport utilization and costs. From this, different options can be derived and presented to the consumer, who then has the opportunity to decide whether to accept a more environmentally friendly option or prefer to have their original delivery requirements fully met.

5 Conclusions, Limitations, and Outlook

The purpose of this paper was to show how consumer choices in regards to delivery options of online purchased groceries affect the sustainability of last mile logistics. Therefore, we developed a simulation model and analyzed two different scenarios. By adapting the baseline scenario to individual consumer preferences, the number of orders is increased, but at the same time, the number of tours is reduced, resulting in a slightly better utilization of vehicle capacity. Even if these improvements are only marginal, the integration of consumer preferences for the delivery of groceries ordered online can be expected to improve rather than worsen the sustainability effects of deliveries. Our results also show that electrically powered vehicles have a decisive influence on the sustainability of delivery. For example, the emissions in the scenario with individual consumer preferences and the exclusive use of electric vehicles are 1.3 kg CO_2eq, while this value is 4.24 kg CO_2eq for the exclusive use of combustion vehicles. Our findings further indicate that vehicle utilization is low in both scenarios. In the baseline scenario, vehicle utilization is around 10%, and in the scenario with empirical data, it is around 15%. Since vehicle utilization is another critical factor in assessing sustainability, there is still significant potential here to make the delivery of groceries ordered online more environmentally friendly.

Even though the costs are a critical factor in the design of the last mile, we have not yet addressed this aspect. This is because the focus of this study is on sustainable and individualized delivery options, and costs, in the long run, will be integrated into our approach as a result of a selected delivery option. For future research, it is, therefore, necessary to assign a possible price to the individual delivery options. Since the delivery options will differ depending on the selection of a delivery option, a dynamic pricing model would be appropriate. As individual adjustments to delivery lead to additional costs for logistics service providers or retailers, the extent to which consumers are willing to contribute to the delivery costs incurred must further be investigated (Witten and Schmidt 2019). We expect that the willingness-to-pay a surcharge for delivery will be higher if the chosen delivery options have a positive impact on sustainability.

For this study, only three features for describing consumer behavior on last mile logistics options were investigated. Nevertheless, it can be seen that there is a direct link between chosen logistics options and responding emissions. Therefore, in future work, sensitive analysis on all the other features needs to be done to investigate the interaction of all features in different circumstances. In the long term, the simulations presented can be integrated into the online ordering process of groceries. After a consumer selects a delivery option for the groceries ordered online, the consumer would get direct feedback on the sustainable impact of that delivery option. This allows consumers to reflect and evaluate their decision. If desired, the consumer can then adjust his decision and tailor the last mile logistics to the individual and sustainable preferences (Freitag and Kotzab 2020).

References

Agatz, N., Campbell, A.M., Fleischmann, M., Savels, M.: Challenges and opportunities in attended home delivery. In: Golden, B., Raghavan, S., Wasil, E. (eds.) The Vehicle Routing Problem: Latest Advances and New Challenges, vol. 43, pp. 379–396. Springer, Boston, MA (2008)

Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Eds.), Handbuch Logistik, VDI-Buch. Springer Berlin Heidelberg, Berlin, Heidelberg. (2008). https://doi.org/10.1007/978-3-540-72929-7

Brabänder, C.: Die Letzte Meile: Definition, Prozess, Kostenrechnung und Gestaltungsfelder (2020). Available at: https://doi.org/10.1007/978-3-658-29927-9 (accessed 29 June 2020)

Caggiani, L., Colovic, A., Prencipe, L.P., Ottomanelli, M.: A green logistics solution for last-mile deliveries considering e-vans and e-cargo bikes. Transportation Research Procedia. **52**, 75–82 (2021)

Clausen, U., Stütz, S., Bernsmann, A., Heinrichmeyer, H.: Die Letzte Meile. ZF-Zukunftsstudie 2016. Fraunhofer IML, Stuttgart, pp. 1–34 (2016)

Deges, F., Speckmann, A.-S.: Lieferservice Im Online-Lebensmittelhandel: Analyse Des Spannungsfeldes Zwischen Den Erwartungen Der Konsumenten Und Den Leistungsversprechen Der Anbieter, Diskussionsbeitrag No. 19, Europäische Fachhochschule Brühl, Brühl (2020)

Dworak, T., Burdick, B.: Ökologische Und Soziale Chancen Und Risiken Des E-Commerce Im Nahrungsmittelsektor. Wuppertal Papers No. 126, Wuppertal Institut für Klima, Umwelt, Energie, Wuppertal (2002). Available at: http://nbn-resolving.de/urn:nbn:de:bsz:wup4-opus-15186

Ehrler, V.C., Schöder, D., Seidel, S.: Challenges and perspectives for the use of electric vehicles for last mile logistics of grocery e-commerce – Findings from case studies in Germany. Res. Transp. Econ. p. 100757 (2019)

Freitag, M., Kotzab, H.: A concept for a consumer-centered sustainable last mile logistics. In: Freitag, M., Haasis, H.-D., Kotzab, H., Pannek, J. (eds.) Dynamics in Logistics, pp. 196–203. Springer, Cham (2020)

Gausmann, O.: Kundenindividuelle Wertschöpfungsnetze, 1. Aufl., Gabler, Wiesbaden (2009). https://doi.org/10.1007/978-3-8349-9917-7

Gevaers, R., Van de Voorde, E., Vanelslander, T.: Characteristics and typology of last mile logistics from an innovation perspective in an urban context. In: Macharis, C., Melo, S. (eds.) City Distribution and Urban Freight Transport: Multiple Perspectives, pp. 56–74. Edward Elgar, Cheltenham (2011)

Grant, D.B., Fernie, J., Schulz, B.: Enablers and barriers in german online food retailing. Suppl Chain Forum Int J. **15**(3), 4–11 (2014)

Hagen, T., Scheel-Kopeinig, S.: Would customers be willing to use an alternative (chargeable) delivery concept for the last mile? Working Papers, No. 17. Res. Transp. Bus. Manag. p. 100626 (2021)

Hagenmeyer, M.: Lebensmittel-Kennzeichnungsverordnung: mit Los-Kennzeichnungs-Verordnung und Fertigpackungsverordnung (Auszug); Kommentar, 2. Aufl., Beck, München (2006)

HDE Online-Monitor.: Handelsverband Deutschalnd, IFH Köln, pp. 1–59 (2021)

Heinemann, G.: Der neue. Online-Handel: Geschäftsmodelle, Geschäftssysteme und Benchmarks im E-Commerce. Springer, Wiesbaden (2020). https://doi.org/10.1007/978-3-658-28204-2

Kläver, K., Profeta, A., Kircher, C. (2020): Kauf-Und Ernährungsverhalten Der Deutschen Verbraucher Vor Und Während Der Covid-19-Pandemie: Tabellen-Und Grafikband, Landesinitiative Ernährungswirtschaft Niedersachsen und Deutes Institut für Lebensmitteltechnik e.V. Available at: https://www.dil-ev.de/fileadmin/user_upload/Covid_19_Studie__1_.pdf

Kölmel, B., Pfefferle, T., Bulander, R.: Mega-Trend Individualisierung: Personalisierte Produkte und Dienstleistungen am Beispiel der Verpackungsbranche. In: Deutscher Dialogmarketing Verband e (Ed.), Dialogmarketing Perspektiven 2018/2019, Springer, Wiesbaden, pp. 243–260 (2019)

Kämäräinen, V., Punakivi, M. (2002). Developing cost-effective operations for the e-grocery supply chain. International Journal of Logistics: Research and Applications, **5**(3), 285–298.

Mangiaracina, R., Perego, A., Seghezzi, A., Tumino, A.: Innovative solutions to increase last-mile delivery efficiency in B2C e-commerce: a literature review. International Journal of Physical Distribution & Logistics Management. **49**(9), 901–920 (2019)

Olivera, A.C., García-Nieto, J.M., Alba, E.: Reducing vehicle emissions and fuel consumption in the city by using particle swarm optimization. Appl. Intell. **42**(3), 389–405 (2015)

Ombuki, B., Ross, B.J., Hanshar, F.: Multi-objective genetic algorithms for vehicle routing problem with time windows. Appl. Intell. **24**(1), 17–30 (2006)

Özarık, S.S., Veelenturf, L.P., van Woensel, T., Laporte, G.: Optimizing e-commerce last-mile vehicle routing and scheduling under uncertain customer presence. Trans. Res. Part E Logist. Transp. Rev. **148**, 102263 (2021)

Pronello, C., Camusso, C., Valentina, R.: Last mile freight distribution and transport operators' needs: Which targets and challenges? Transp. Res. Proc. **25**, 888–899 (2017)

Reichert, D., Cito, C., Barjasic, I.: Ressourceneffizienz und Nachhaltigkeit. In: Lean & Green: Best Practice, pp. 5–23. Springer, Wiesbaden (2018)

Reichwald, R., Piller, F., Ihl, C.: Interaktive Wertschöpfung: Open Innovation, Individualisierung und neue Formen der Arbeitsteilung. Gabler, Wiesbaden (2009)

Saenz, J., Figliozzi, M., Faulin, J.: Assessment of the Carbon Footprint Reductions of Tricycle Logistics Services. Transp. Res. Record J. Transp. Res. Board. **2570**(1), 48–56 (2016)

Savelsbergh, M., Van Woensel, T.: 50th Anniversary Invited Article—City Logistics: Challenges and Opportunities. Transp. Sci. **50**(2), 579–590 (2016)

Schewelsky, M., Follmer, R. Dickmann, C.: CO2-Fußabdrücke Im Alltagsverkehr: Datenauswertung Auf Basis Der Studie Mobilität in Deutschland, edited by Umweltbundesamt (2020)

Schnedlitz, P., Lienbacher, E., Waldegg-Lindl, B., Waldegg-Lindl, M.: Last Mile: Die letzten – und teuersten – Meter zum Kunden im B2C ECommerce. In: Crockford, G., Ritschel, F., Schmieder, U.-M. (eds.) Handel in Theorie und Praxis, pp. 249–273. Springer, Wiesbaden (2013)

Seidel, S., Mareï, N., Blanquart, C.: Innovations in e-grocery and Logistics Solutions for Cities. Transp. Res. Proc. **12**, 825–835 (2016)

Shaik, A.M., Rao, V.V.S.K., Rao, C.S.: Development of modular manufacturing systems—a review. Int. J. Adv. Manuf. Technol. **76**(5–8), 789–802 (2015)

Shen, L., Tao, F. and Wang, S.: Multi-Depot Open Vehicle Routing Problem with Time Windows Based on Carbon Trading. Int. J. Environ. Res. Publ. Health 15(9) (2018). https://doi.org/10.3390/ijerph15092025

State Statistical Office Bremen (Ed.). Population and Employment, Households and Families in the State of Bremen 2011 to 2015: Results of the Microcensus with Time Series 2005 to 2015 (2016). Available at: https://www.statistik.bremen.de/sixcms/media.php/13/AI_AVI_MZ_2011-2015_pdfa.pdf

Ternès, A., Towers, I., Jerusel, M.: Konsumentenverhalten im Zeitalter der Mass Customization: Trends; Individualisierung und Nachhaltigkeit. Springer, Wiesbaden (2015)

Umundum, P.: Die letzte Meile – Königsdisziplin der Logistik. In: Voß, P.H. (ed.) Logistik – die unterschätzte Zukunftsindustrie, pp. 149–162. Springer, Wiesbaden (2020)

Vahrenkamp, R., Kotzab, H., Siepermann, C.: Logistik: Management und Strategien, 7., überarbeitete und erweiterte Auflage. Oldenbourg, München (2012)

Wegner, K.: Potenziale der Digitalisierung für die letzte Meile in der Logistik. In: Schröder, M., Wegner, K. (eds.) Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains, pp. 285–301. Springer, Wiesbaden (2019)

Winkenbach, M., Janjevic, M.: Classification of Last-Mile Delivery Models for e-Commerce Distribution: A Global Perspective. In: Taniguchi, E., Thompson, R.G. (eds.) City Logistics 1, pp. 209–229. Wiley, Hoboken, NJ (2018)

Witten, P., Schmidt, C.: Globale Trends und die Konsequenzen für die Logistik der letzten Meile. In: Schröder, M., Wegner, K. (eds.) Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains, pp. 303–319. Springer, Wiesbaden (2019)

Yumurtacı Hüseyinoğlu, I.Ö., Kotzab, H., Köstepen, K.G., Halaszovich, T.: Assessing consumer logistics functions in grocery shopping: Evidence from an emerging market. J. Mark. Channels. **26**(1), 72–86 (2020)

Logistics Challenges Along the New Silk Roads

Hans-Dietrich Haasis, Jianhui Du, and Xuejun Sun

Abstract In 2013, Chinese President Xi Jinping advised to establish the "Silk Road Economic Belt" and the "21st Century Maritime Silk Road," also referred as the Belt and Road Initiative or the New Silk Roads Policy. The intention is to promote international and regional trade as well as cooperation in and between Asia and Europe. Consequently, international maritime and terrestrial freight transport corridors are either established or strengthened and operated. The purpose of this paper is to reflect the Belt and Road Initiative from the perspective of logistics. The aim is to identify and formulate circumstances, expectations, opportunities, and peculiarities of logistics along the New Silk Roads. For this purpose, four corresponding challenges will be considered and outlined after an introduction to the Belt and Road Initiative. The four logistics challenges concern the awareness of new freight transport corridors and the assessment of possibilities for opening new transport relations and new markets, the implementation of new and the adaptation of existing supply chains to increase strategic logistics flexibility, the availability and use of digital infrastructure and connectivity for improved communication and coordination of logistical processes, and the willingness to consider regional and cultural differences in the preparation and realization of supply chain decisions.

1 New Silk Roads Policy and Its Implications

In 2013, China proposed to establish the "Silk Road Economic Belt" and the "21st Century Maritime Silk Road," also referred to the long-term development policy called Belt and Road Initiative, One Belt One Road, or New Silk Roads. Chinese President Xi Jinping first explains his vision of revitalizing the ancient Silk

---

H.-D. Haasis (✉)
Faculty of Business Studies and Economics, University of Bremen, Bremen, Germany
e-mail: haasis@uni-bremen.de

J. Du · X. Sun
Zhongyuan University of Technology, Zhengzhou, Henan, China
e-mail: jhd@zut.edu.cn; sunriver@126.com

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_14

Road to jointly build the Silk Road Economic Belt at Nazarbayev University on September 7, 2013, as part of his state visit to Kazakhstan. One month later, on October 3, 2013, during his state visit to Indonesia he announced the 21st Century Maritime Silk Road. The Silk Road Economic Belt represents the land-based trade corridors linking Asia and Europe by establishing roads, railways, and pipelines. The 21st Century Maritime Silk Road indicates the maritime shipping passages spanning across the South China Sea, the South Pacific Ocean, and the Indian Ocean area. China's intention is to promote regional trade and cooperation through improving connectivity in Asia, Africa, and Europe. Chinese authorities encourage domestic enterprises to go global and actively cooperate with other states. Trade liberalization agreements are signed between China and countries influenced by the Belt and Road Initiative to promote trade and investment environments. Additional information is provided, for example, within (Lam et al. 2018; Lee et al. 2018; National Development and Reform Commission 2015; Wan 2021; Ye 2020; Ye and Haasis 2018).

In connection with infrastructure construction projects, large amounts of Chinese investments are already realized and will be realized in the future. Examples are, amongst others, the Port of Piraeus in Greece as well as the new Eurasian railway link connecting China, Central Asia and Europe. The Chinese National Development and Reform Commission, Ministry of Foreign Affairs and Ministry of Commerce, already in March 2015 released the first statement "Vision and Actions on Jointly Building Silk Road Economic Belt and 21st-Century Maritime Silk Road," which illustrates a grand blueprint as the strategic-level framework (National Development and Reform Commission 2015). The statement presents five major goals to enhance regional cooperation: policy coordination, facilities connectivity, unimpeded trade, financial integration, and people-to-people bonds. In the same year, Chinese government set up the Office of the Leading Group for Promoting the Belt and Road Initiative.

Although there are threats related to the cooperation arrangements like financial investment risks and political influence, the chances may outweigh according to the economic strengths and technological weaknesses of the countries and companies engaged in. At present, there are more than 50 countries signed cooperation agreements with China, according to the Chinese "Belt and Road Portal" (https://eng.yidaiyilu.gov.cn/, last accessed 2021/06/25).

The pattern of logistics networks and supply chains is likely to be reshaped, and more countries and regions as well as their markets may become favorable for international enterprises and motivate facilities relocation processes (Hammami and Frein 2014; Khan and Haasis 2020; Rodrigue and Hesse 2006; Ye 2020; Ye and Haasis 2018). Against this background, logistics challenges may occur, for example, the revision of supply chain designs with respect to the availability of freight transport corridors, the ensuring of strategic logistics flexibility considering new logistics hubs and freight villages along the New Silk Roads, the accessibility to digital processes and communication platforms reducing digital divide conflicts, as well as the significance and appreciation of cultural influences on interregional supply chain decisions.

2 Freight Transport Corridors and Supply Chain Design

Efficient as well as well-developed freight transport corridors are important, both for Europe and China as well as for the international networking of essential production sites with procurement and sales markets (Hammami and Frein 2014; Lee et al. 2018; Rodrigue and Hesse 2006). Their design and expansion aim

- to ensure economic accessibility for supply chains,
- to increase the attractiveness and visibility of the regions along the transport corridors,
- to increase the reliability, the cost-effectiveness, and the safety of freight transportation, as well as
- to ensure a future-oriented sustainable development of economic regions.

If this succeeds, transport corridors are valuable enablers for transport and trade facilitation, and thus for economic and social prosperity.

The demand for long-distance border-crossing freight transport is continuously growing in the economic globalization era, which puts a lot of pressure on the current global freight transportation system as well as on supply chains. Some transportation geographers also studied the transportation infrastructure and networks, container terminals, gateways, ports cooperation and regionalization, transportation network, etc., for example, see (Dovbischuk and Haasis 2011; Rodrigue 2012; Rodrigue and Hesse 2006).

Against this background, the terrestrial railway connections and the maritime transport relations to and from China play an essential role at the international level. This is mainly true in relation to the Chinese Belt and Road Initiative (Lam et al. 2018; Lee et al. 2018; National Development and Reform Commission 2015; Ye 2020; Ye and Haasis 2018). Figure 1 outlines the infrastructure network focusing

Fig. 1 Belt and Road Initiative—The infrastructure network (Source: Mercator Institute for China Studies, Berlin, 2018)

on the Belt and Road Initiative. In addition, for Europe, the North-South axes as well as the West-East axes play an important role, i.e., the connection between Scandinavia and the Iberian Peninsula, the connection between the Baltic Sea and the Black Sea as well as the connection between Rotterdam, Duisburg, Warsaw, and Moscow. Focusing on Germany, the seaport hinterland connections are crucial for the performance of the German economy.

Congestion in important ports and insufficient capacity of distribution channels become bottlenecks for development and for the operation of supply chains. China's Belt and Road Initiative is an infrastructure-led development policy with large amounts of investments flowing into transportation infrastructure projects. These efforts are expected to expand the capacities of some critical ports and increase the connectivity of specific countries and regions to the global transportation system. Transportation infrastructure will affect firms' strategies about their network configuration and their supply chain design (Buer et al. 2019; Khan and Haasis 2020; Lee et al. 2018; Rodrigue and Hesse 2006; Ye 2020; Ye and Haasis 2018). The port of Piraeus as a case under the frame of the Belt and Road Initiative is a good example to illustrate how transportation infrastructure improvement affects the importance of this node in the entire network (Ye and Haasis 2018).

Of course, also the China Railway Express, operated by the state-owned enterprise China Railway Corporation, enables international container freight transport services between China and Europe. In the "Development Plan of China-Europe Freight Train Construction (2016–20)," issued by the Leading Group Office on the Construction of the Belt and Road (please see: https://eng.yidaiyilu.gov.cn/, last accessed 2021/06/25), the background and the demand for railway freight transport between China and European countries are analyzed. Moreover, the spatial structure design of the railway transit corridors, hub nodes, and routes is analyzed and published (Ye 2020).

There are three main railway passages: The western routes have three main branches; one runs through Kazakhstan connecting to the Trans-Siberian Railway; one crosses the border at Horgos Border port; one runs across Torugart pass and connects to the planned China-Kyrgyzstan-Uzbekistan Railway. The central route runs across Erlianhot land port and through Mongolia to connect to the Trans-Siberian Railway. The Eastern route goes across Manzhouli land port linking to Trans-Siberian Railway. All these passages end in European destinations.

Recently, enterprises are increasingly interested in relocating facilities; network redesign processes have become more frequent. Of course, enterprises are motivated by varying reasons such as offshoring, expansion opportunities to new markets, mergers and acquisitions, financial and tax advantages offered by countries (Hammami and Frein 2014).

In the context of the Belt and Road Initiative, the connectivity of selected locations is improved through investments in transportation infrastructure. Thus, national and regional governments should take advantage of these improvements and offer a more favorable business environment to attract enterprises' investments. On the other side, enterprises may reconsider their current network configuration and adapt their supply chain strategies to respond to the possible opportunities induced by the implementation of the Belt and Road Initiative (Ye 2020; Ye and Haasis 2018).

For the design and operation of these freight transport connections, it is important to consider the challenges of today and of the future, in particular the demands on freight mobility, the availability of energy resources, the efforts to protect the climate, the changes caused by the digital transformation of processes, and the need to integrate transport systems. Particularly from the point of view of climate and resources, rail transport and maritime transport play an important role (Tran et al. 2017). Finally, in view of the importance for the future freight transport system, investments in the provision and use of information and knowledge along the transport connections must also be considered (National Development and Reform Commission 2015; Wan 2021).

This shows that when dealing with supply chains along international or intercontinental freight transport routes, it is not only necessary to contribute to spatial planning and construction industry, but also to invest more than ever in the provision and use of information as well as to invest in the development and the assessment of innovative cooperation and business models.

3 Logistics Hubs, Freight Villages, and Strategic Logistics Flexibility

Considering the availability of new freight transport corridors, the strategic flexibility of an enterprise related to the operation of supply chains may increase. According to the supply chain design as well as the supply chain operation, these new options for freight transportation give the chance and supports to better respond to changes in the supply network access and the availability of logistics services along the long-distance transportation. The strategic logistics flexibility may be strengthened and improved. However, this depends on information and the presence and the scope of business activities in relation to logistics hubs and freight villages along the New Silk Roads.

Freight villages are specific logistics hubs. According to the DGG—Deutsche GVZ-Gesellschaft mbH (please see: gvz-org.de, last accessed 2021/06/25), they are defined as “logistics centers, where the cargo from different transport modes can be reloaded, compiled, and prepared for transportation. That place links and brings together different transport modes like road and rail, transport companies, supplementary transport services as well as industrial and trading companies. The spatial proximity promotes coopetition and division of labor of the enterprises on site. A major function of the freight villages is the management of politically promoted combined transport and the shifting of cargo traffic from road to rail. The cooperation of the companies on site makes it possible to realize high capacity utilization.”

More and more logistics hubs and freight villages like this can be found along the maritime shipping routes and, of course, along the three main railway passages between China and Europe, for example, the Horgos border port in Xinjiang, the Erlianhot land port and the Manzhouli land port.

As a transportation interface and a business area for logistics-focused services, freight villages are gaining wide acceptance in countries worldwide. Nowadays, the style on how to operate and collaborate in freight villages is often pointed out as a way for green logistics and sustainable freight transportation (Wu 2013; Wu and Haasis 2018). On the other hand, they provide added value services, such as inventory management, high-density warehousing, and packaging on behalf of manufacturing, retailing, and wholesaling customers. In addition, a freight village provides auxiliary facilities such as warehouses, groupage systems, customs, maintenance workshops, banks, insurance offices. Maybe in the future, they also provide decentralized additive manufacturing by using 3-dimensional printers (Barz et al. 2016), for example, in connection with customization of products for the customer.

Thus, freight villages are the foundation for intelligent and integrated freight transport and supply chain management on a regional, a national, and an international scale.

Logistics hubs and freight villages must be distinguished from so-called Special Economic Zones (Aggarwal 2011; Khan and Haasis 2020). To attract foreign direct investment these zones can be found within or nearby logistics hubs. However, the concept is focused on financial and tax policies. Whereas the aim of freight villages is to improve freight transportation and to motivate collaboration between coopetitive enterprises for a more efficient, green, and reliable supply chain.

On a regional scale, freight villages may be realized to organize regional and urban freight transport to and from sales markets or cities. On a national scale, freight villages support efficient, green, and reliable freight distribution between national supplier and sales markets using electricity-based rail transportation. And on an international scale, freight villages open the door to international transportation networks and markets. Thus, there are advantages for all stakeholders interested in transportation, for example, the customers, the workers, the regional authorities, the railway operators, the logistics service providers, the suppliers and the industrial companies by increasing their economic and social development and welfare.

Challenges linked to logistics hubs and freight villages are related to reliable connectivity, the organizational structure, the coopetitive business model, and the use of knowledge management for the improvement of collaboration. Investing in New Silk Roads also means to invest in capabilities, knowledge, and training.

Nowadays, the concept of logistics hubs and freight villages furthermore can be explained by the new economic geography theory, where spatial and agglomeration economies are the focal point. According to this theory, agglomeration is attributed to increasing returns of scales based on knowledge spillovers, market demand and cost of trade, as well as to infrastructure quality, the collaboration of enterprises and better technical and business environment (Aggarwal 2011; Dovbischuk and Haasis 2011).

4 Digital Silk Roads, Platform Ecosystems, and Digital Divide

Since several years the topic of digitalization has been discussed worldwide in connection with a better operation of processes as well as a shift of process control either or both, to the objects and the international web. This applies mainly to logistical processes (Freitag et al. 2020; Haasis et al. 2015; Jahn et al. 2020; Kersten et al. 2020; Wan 2021). In this context, digital transformation is the process of implementing digital technologies and supporting capabilities to create digital business models. However, due to high transaction costs in combination with a revision of business development and an implementation of change-management actions, the progress in digitalization goes on step by step and is obviously different between enterprises, regions, and states according to their related technological and human capabilities. The implementation of a digital silk road may support to overcome the dilemma of the digital divide.

Nevertheless, a rethinking of communication, coordination, and cooperation between stakeholders along a supply chain is obvious. Mainly concerning information management between partners in supply chains, there are substantial potentials for improving efficiencies. For example, efficiency improvements can be obtained by a better coordination along the international supply chain, by a more comprehensive and coopetitive regional cooperation, as well as by an innovative shifting of logistics control mechanisms into the cloud. By far, the options of modern information and communication technologies are not used sufficiently.

More and more decision-makers on a business level as well as on a political level focus on the design and operation of digital web-based service platforms based on the landlord philosophy. Public institutions may provide the digital infrastructure, and data transfer interfaces should be standardized across enterprises. The digital suprastructure could be developed and provided by private enterprises. Part of the digital suprastructure could be applications for the use of on-demand services, the use of logistics-oriented cyber-physical systems, the use of methods for data mining and data analyzing as well as the use of other coopetitive decision logics.

Thus, research and development towards collaborative decision-making, based on a coopetitive corporate policy, in ports, freight villages and in supply chains should be pushed by research institutions in close cooperation with enterprises. The same holds for the development of new data-driven control models for supply chain management and logistics operations, as well as for new business models (Haasis et al. 2015). In connection with business models, incentive schemes may be developed, overcoming the political motivated regional competition and realizing a coopetitive multi-plant cost calculation and a multi-plant profit allocation based on transfer payments. Of course, the incentive schemes must consider legal and operational requirements. Regarding new business models, the allocation of processes to process owners must be reconsidered. In the future, process owners of today may be substituted by process owners of tomorrow.

As known, in general, seaports are competitors, both from a business as well as from a regional political point of view. Of course, this is also true for terminal operators, logistics service providers and forwarders. Nevertheless, considering bottleneck situations, for example, stressed by port operations and by hinterland transportation as well as focusing on investments for port development and infrastructure provision, it may be an option to cooperate from time to time. This may lead to advantages like less waiting times in front of gates and terminals, better usage of straddle carriers and container yard capacities, as well as more time-efficient transshipment processes. Due to this, seaport communication and cooperation reflect typical situations for coopetition. The economic success of this coopetition may be enabled by an improved supply of digital information used for better communication and coordination between partners in the maritime supply chain as well as by a better allocation of scarce resources in line with collaborative decision-making. Examples for digital service innovations based on collaborative decision-making are yard and container lot management, ship announcement after re-direction and terminal preparation, gate drive-through management as well as a flexible and conditional supply chain management.

Maybe in the future a container ship decides autonomously about the best terminal and berth for container transshipment, or a vessel moored at a quay controls independently the terminal operations for its containers. Or a retailer will control the transshipment of requested containers towards the hinterland transportation to the demanding shop based on the specified demand of customer products. Or the container itself self-controls some of his operations on the container yard. These cloud-based applications in the future may be operated by using a digital web-based service platform across locations, upgraded on the base of port community systems and of terminal operating systems. The platform could have the characteristics of an essential facility. The related process model could be developed further towards a port-as-a-service-model. The port can be called a smart service port (Jahn et al. 2020).

Automation is already a real up-to-date development, for example, in selected terminals in the seaports of Hamburg, Busan, and Dubai. Considering the Next Generation Port 2030 in Singapore or the well-known TradeLens initiative of Maersk and IBM based on a blockchain-enabled digital shipping platform, the seaport 4.0 is approaching. Based on the philosophy of the well-known landlord model, as already mentioned, the digital platform could be provided as standardized public infrastructure, and the service applications could reflect the customer-orientated private suprastructures. Algorithms and new data-driven applications for berth allocation, container yard operations, storage management and crane scheduling policies can be shifted towards cloud solutions. Of course, similar considerations can be made for freight villages and, in general, for logistics hubs.

Thus, in the future the services could be offered by new Internet enterprises powered by data sources. The questions in the future are who will provide the digital business models on demand, who has the data power and on how to handle the digital divide between companies and regions. Therefore, related market consequences for seaports, freight villages and supply chains must be observed, analyzed, and challenged.

Finally, we must decide about the extent of external web-based control. This is a thrilling and challenging topic that should be covered in the context of discussions about the ethics of logistics.

The idea to achieve digital connectivity by integrating the concept of the internet of things and e-commerce into the Belt and Road Initiative was already mentioned in a White Paper prepared by the National Development and Reform Commission in 2015. In addition, the cross-border laying of optical cables and the expansion of communication networks were called for (National Development and Reform Commission 2015).

This policy of Digital Silk Roads was underlined by Lu Wei, the Director of the so-called Cyberspace Administration, in the China-EU Roundtable for Digital Cooperation in 2015. He mentioned: “We can build a digital silk road, a silk road in cyberspace” (Brown 2017).

5 Interregional Supply Chain Decisions and Cultural Influences

The implementation and the operation of supply chains along the New Silk Roads must deal with decision-makers and logistics service people with different cultural backgrounds. This is of course very clear in the widely recognized and accepted consideration of public holidays. However, these backgrounds may have an influence on the art on how to decide and on how to operate efficient and reliable logistics processes (Baumann et al. 2013; Fawcett et al. 2004; Ferraro and Briedy 2017; Hofstede et al. 2010; Smyrlis 2004). On the other hand, these backgrounds may result in the transfer of ideas, experiences, and knowledge between the regions and people along the New Silk Roads, as already happened in the past. Thus, besides only focusing on the supply chains of freight and goods, it is enhancing and worthwhile also to focus on the cultural chain of ideas and knowledge to increase the understanding and appreciation of decision-makers and people communicating, cooperating, and working together.

Maybe the focus on the people and their cultural differences, and by this on the cultural chain, could support the improvement of the logistics performance of companies and regions. Still, in some parts of Asia, contracts are just a general commitment to do business together and are less meaningful than personal relationships between individuals. Or what you believe to be a standard performance contract with a warehouse service provider containing a “circumstances change” clause, which essentially means the contract is not legally binding (Smyrlis 2004). According to the Logistics Performance Index, a benchmarking tool published by the World Bank (Arvis et al. 2018), there are still numerous challenges, but also opportunities, along the New Silk Roads, influencing the logistics performance, which must be handled by logistics people with different cultural backgrounds. The indicators considered for calculating the Logistics Performance Index are related to customs, infrastructure, service quality, timeliness, international shipping, and tracking and tracing. The characteristics of these indicators depend on people's decisions. By the way, the latest edition of the Logistics Performance Index allows for comparisons across 160 countries (Arvis et al. 2018): In 2018, Germany achieved rank 1, China was on rank 27, Kazakhstan on rank 71, and the Russian Federation on rank 75.

Obviously, the problems to ensure time-efficient and reliable processes are not only different languages, time zones, and measurement units, but behavior patterns and context interpretations. In connection with implementing logistics information systems and decision support tools, it is not uncommon, mainly in developing countries, to hear something like the following: "We were used to work without pressure and postponing some tasks for later because the system was not synchronized. Everyone was working for themselves without pressure. However, with this synchronized system, work must be always done instantaneously so that operations can take place on time. As a result, we feel that we are under pressure from the system and that we are no longer acting according to our own free will, but that we are subject to the system." This statement may underline those cultural aspects must be considered as part of interregional supply chain decisions to avoid misunderstanding and miscommunication. As well known, miscommunication increases uncertainty and risk and thus costs and can lead to wrong supply chain decisions.

This perspective clearly shows the importance of knowledge management within a region, a company, as well as between companies working together (Wu 2013). Human-centric knowledge management may help to identify cultural differences and to discuss these differences within an intercultural team providing new solutions for better cooperation and supply chain management. From this point of view, corresponding fresh reference should be given to topics like business process modeling, re-inventing of business models, business data analytics as well as the use of artificial intelligence for decision support and negotiation-orientated decision models (Baumann et al. 2013; Haasis et al. 2015).

6 Synopsis and Conclusions

Within this paper**,** logistics challenges are outlined related to the realization and operation of modified as well as new maritime and terrestrial silk roads based on the Belt and Road Initiative. Of course, besides the available infrastructure, people's decisions, services, business informatics, data analytics and innovation mainly influence the success of logistics both on a regional level as well as on an international level. Knowing that world trade flows are interconnected and changeable, logistics must nevertheless ensure distributed production and supply in a dynamic environment.

The aspects discussed are based on literature reviews as well as on studies accomplished at the Chair of Maritime Business and Logistics in Bremen in cooperation with researchers from China. The challenges mentioned within four chapters relate to the awareness of new freight transport corridors and the assessment of possibilities for opening new transport relations and new markets, the implementation of new and the adaptation of existing supply chains to increase strategic logistics flexibility, the availability and use of digital infrastructure and connectivity for improved communication and coordination of logistical processes, and the willingness to consider regional and cultural differences in the preparation and realization of supply chain decisions.

Of course, transport corridors are valuable enablers for transport and trade facilitation. However, decision-maker should be aware of the new corridors and should be able to evaluate both the supply chain risk as well as the economic and sustainable supply chain potential.

Strategic logistics flexibility may be increased by considering new logistics hubs and freight villages. However, this depends on information as well as on the presence and the scope of business activities in relation to logistics hubs and freight villages along the New Silk Roads. Thus, logistics hubs and freight villages are the foundation for intelligent and integrated freight transport as well as for supply chain management on a regional, a national, and an international scale.

The access to digital processes, innovations, and communication platforms is essential to increase the efficiency of logistics processes and supply chains. However, this access varies from company to company and from country to country. Moreover, the questions in the future are who will provide the digital business models and who has the data power. The implementation of a digital silk road may support to overcome at least the dilemma of the digital divide.

The operation of supply chains along the New Silk Roads takes into account decision-makers and logistics service people with different cultural backgrounds. Maybe a stronger focus on the people and their cultural differences, and by this on the cultural chain, could support the improvement of the logistics performance of companies and regions.

These remarks may motivate not only to invest in infrastructure, i.e., in ports, railway connections and warehouses, but also to invest in knowledge, research, education, training, communication, and understanding. By this, logistics actions along the New Silk Roads may serve as enabler and promoter for sustainable welfare as well as for a better international communication and peaceful cooperation from the point of view of states, regions, enterprises, and people.

Acknowledgments This paper is based on a project funded by the German Academic Exchange Service (DAAD) with financial resources from the Federal Ministry for Economic Cooperation and Development.

References

Aggarwal, A.: Promoting agglomeration economies and industrial clustering through SEZs. J. Int. Econ. Commer. Policy. **2**(2), 201–227 (2011)

Arvis, J.-F., Ojala, L., et al. (eds.): Connecting to Compete. Trade Logistics in the Global Economy. The World Bank, Washington (2018)

Barz, A., Buer, T., Haasis, H.-D.: Quantifying the effects of additive manufacturing on supply networks by means of a facility location-allocation model. Logistics Research 9(1), 13, pp. 1–14 (2016)

Baumann, T., Haasis, H.-D., Nehlsen-Pein, T.: Human decision making in business. Implications and application operations of neurosciences for business decisions. Bus. Syst. Rev. **2**(1), 1–14 (2013)

Brown, R.: Beijing's Silk Road goes digital, www.cfr.org/blog/beijings-silk-road-goes-digital, last accessed 2021/06/25 (2017)

Buer, T., Haasis, H.-D., Kinra, A., Kotzab, H.: An overview to contemporary maritime logistics and supply chain management decision areas. In: Panayides, P.M. (ed.) The Routledge Handbook of Maritime Management, pp. 113–123. Routledge, London (2019)

Dovbischuk, I., Haasis, H.-D.: Path dependency in the process of sustainable regional development of transport and logistics regions. In: Ivanov, D., Kopfer, H., Haasis, H.-D., Schönberger, J. (eds.) Dynamics and Sustainability in International Logistics and Supply Chain Management, pp. 241–247. Cuvillier, Göttingen (2011)

Fawcett, S.E., Magnan, G.M., Williams, A.J.: Supply chain trust is within your grasp. Suppl. Chain Manag. Rev. **8**(2), 20–26 (2004)

Ferraro, G.P., Briedy, E.K.: The cultural dimension of global business, 8th edn. Routledge, London (2017)

Freitag, M., Haasis, H.-D., Kotzab, H., Pannek, J. (eds.): Dynamics in Logistics. Springer, Cham (2020)

Haasis, H.-D., Landwehr, T., Kille, G., Obsadny, M.: Cloud-based eBusiness standardization in the maritime supply chain. In: Dethloff, J., Haasis, H.-D., Kopfer, H., Kotzab, H., Schönberger, J. (eds.) Logistics Management. Products, Actors, Technologies, pp. 256–276. Springer, Berlin (2015)

Hammami, R., Frein, Y.: Redesign of global supply chains with integration of transfer pricing: Mathematical modeling and managerial insights. Int. J. Prod. Econ. **158**, 267–277 (2014)

Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organisations: Software of the Mind, 3rd edn. McGraw-Hill, New York (2010)

Jahn, C., Kersten, W., Ringle, C. M. (eds.) Data science in maritime and city logistics. epubli, Berlin (2020)

Kersten, W., Blecker, T., Ringle, C. M. (eds.) Digitalization in Supply Chain Management and Logistics. epubli, Berlin (2020)

Khan, A., Haasis, H.-D.: Spatially induced effects and sustainability for special economic zones: implications for zones in Pakistan under China Pakistan Economic Corridor. Region. Sci. Inquiry J. **2**(12), 179–194 (2020)

Lam, J., Cullinane, K., Lee, S.-W.: The 21st-century Maritime Silk Road: Challenges and opportunities for transport management and practice. Transp. Rev. **38**(4), 413–415 (2018)

Lee, G., Lee, S.-W., Feng, X.: Special Issue on 'Challenges and chances of the Belt and Road Initiative at the maritime policy and management level'. Marit. Policy Manag. **45**(3), 279–281 (2018)

National Development and Reform Commission (ed.) Vision and actions on jointly building Silk Road Economic Belt and 21st Century Maritime Silk Road, http://webcache.googleusercontent.com/search?q=cache:http://en.ndrc.gov.cn/newsrelease/201503/t201503 30_669367.html, last accessed 2021/06/25 (2015)

Rodrigue, J.-P.: The geography of global supply chains: Evidence from third-party logistics. J. Supply Chain Manag. **48**(3), 15–23 (2012)

Rodrigue, J.-P., Hesse, M.: Global production networks and the role of logistics and transportation. Growth Chang. **37**(4), 499–509 (2006)

Smyrlis, L.: Cultural differences can trump the most logical of supply chain planning. Canadian Transport. Logist. **107**(9), 4–5 (2004)

Tran, N.K., Haasis, H.-D., Buer, T.: Container shipping route design incorporating the costs of shipping, inland/feeder transport, inventory and CO_2 emissions. Marit. Econ. Logist. **19**(4), 667–694 (2017)

Wan, Y: Improving the information flows and business process in global maritime logistics chains considering digitization and digitalization aspects. PhD thesis. University of Bremen, 2021

Wu, J.: Sustainable freight village concepts for agricultural product logistics. A knowledge management-based study. PhD thesis. University of Bremen, 2013

Wu, J., Haasis, H.-D.: The freight village as a pathway to sustainable agricultural products logistics in China. Journal of Cleaner Production, pp. 1227–1238 (2018)

Ye, J.: Regional orientated global logistics networks redesign with respect to the Belt and Road Initiative. PhD thesis. University of Bremen, 2020

Ye, J., Haasis, H.-D.: Investment of the Belt and Road Initiative. In: Freitag, M., Kotzab, H., Pannek, J. (eds.) Dynamics in Logistics, pp. 250–254. Springer, Heidelberg (2018)

Observations and Perceptions on a Doctoral Research Seminar in Engineering and Logistics

Dave A. Louis and Ingrid Rügge

Abstract Logistics is a global challenge needing cooperation across disciplinary as well as cultural diversity. International, interdisciplinary education has the potential to provide meaningful experiences and cultural exchange for the individuals involved, including students, faculty, and university professionals and can have a positive impact on their lives and professionalism. We introduce the doctoral training program International Graduate School for Dynamics in Logistics (IGS) at the University of Bremen and explores a German-American educational partnership on the level of doctoral training. Utilizing scholarly personal narratives, the perceptions of two university professionals were garnered. From the narratives, they discussed the advantages and difficulties associated with cultural exchanges, the exposure to many cultures, and engaging with differing educational systems converging within one program. This qualitative study provides and glimpses into the complexity and intentional nuances that should be addressed when developing and operationalizing international, interdisciplinary education.

1 Introduction

This chapter focuses on an educational scientific examination of the phenomena that internationalization adds to the already challenging claim of interdisciplinarity. In the first part, the history, the concept, and thematic orientation of the interdisciplinary doctoral training group of the Bremen Research Cluster for Dynamics in Logistics of the University of Bremen is presented. Subsequently, an international cooperation with Texas Tech University in the United States and with mostly

---

D. A. Louis
College of Education, University of Houston, Houston, USA
e-mail: dalouis@central.uh.edu

I. Rügge (✉)
International Graduate School for Dynamics in Logistics, University of Bremen, Bremen, Germany
e-mail: ruegge@uni-bremen.de

M. Freitag et al. (eds.), *Dynamics in Logistics*,
https://doi.org/10.1007/978-3-030-88662-2_15

Ethiopian doctoral students is taken as an example, examined with a qualitative approach.

2 Part 1: A Doctoral Training Program for Logistics

The University of Bremen is a very young university that started with alternative study programs and exceptional curricular elements. The founding objectives have been interdisciplinarity, orientation to the practice, and social accountability in teaching and research from the beginning on. Over the years, these guiding principles have been expanded by gender equality, ecological accountability, and internationalization (Universität Bremen 2021). In 1995, several research groups from four different departments of the University of Bremen agreed to build a research cluster on logistics without making it a discipline in its own right. They identified logistics as a multidisciplinary and international area of research, transfer, and education.

Now, the established Bremen Research Cluster for Dynamics in Logistics (Log*Dynamics*) is a cooperative network of research groups from five faculties of the University of Bremen: Production Engineering, Mathematics/Computer Science, Physics/Electrical Engineering, Business Studies/Economics, and Law. Associated partners are the BIBA—Bremer Institut für Produktion und Logistik GmbH, the Institute of Shipping Economics and Logistics (ISL), and the Jacobs University Bremen gGmbH. The fields of activity range from fundamental and applied research to transferring results into practice.

When starting the implementation of the research cluster, the idea arose to support young researchers in this interdisciplinary field in a structured manner. At that time, there was already structured doctoral education in Germany, but very rare in the engineering sciences. With a Master (or the old German degree “Diplom”) German doctoral students usually are not treated as students anymore. They are doctoral candidates: young, independent researchers, working for their living in sciences and self-responsible for any further education. Doctorates of research assistants were common practice, mostly in German and with a low proportion of foreigners.

As Log*Dynamics* recognized the need for systematization of doctoral training and its internationalization, especially in logistics, it set up the International Graduate School for Dynamics in Logistics (IGS) in 2014, doctoral students were recruited. Funding for the IGS began with an initial funding by the state of Bremen, which provided funds for infrastructure, structural supervision, and scholarships. In the meantime, the doctoral scholarships come from the DAAD, from the home countries of the international scientists, or from the Erasmus funding schemes of the EU. The IGS has been a partner in three Erasmus Mundus mobility projects with Asia (cLINK, FUSION, and gLINK) from 2012 to 2018. Furthermore, it hosted twice in a row a PhD SummerCamp for Ethiopian doctoral students from Texas Tech University, USA (TTU Summer Camp 2018). Since 2021, the IGS is one of two

European partners in the Erasmus+ Capacity Building Project SSAPI with Asia. Structural support of the IGS still continues to be funded locally.

The first structuring of the new doctoral training program was taken from already successfully implemented ones in the social sciences, humanities, and marine sciences at the University of Bremen. The IGS adapted it to the conditions in engineering and to the interdisciplinary and international aims of the research cluster. The resulting support system was recognized by all disciplines involved in Log*Dynamics*. The curriculum of the IGS is tested, evaluated, and continuously improved to this day (Rügge and Himstedt 2015; Rügge and Scholz-Reiter 2011; Scholz-Reiter and Rügge 2011).

The incoming scholarship holders pursuing their doctoral research at the IGS benefit from disciplinary supervision, scientific mentoring as well as organizational and social support. Beside the individual doctorate projects, the curriculum covers subject-specific courses, interdisciplinary research colloquiums, dialogue forums with the industry, excursions, international conferences, and individual coaching. The IGS integrates visiting professors into the supervision of the theses and external experts for specific training in the field of personnel development. The working language is English. The objective is to foster excellence in higher education and research by providing an optimal environment in the field of logistics research on different levels.

The IGS meets the challenge of globalization in logistics through doctoral research in the same areas Log*Dynamics* is focusing on. The individual research is centered on four topic areas:

- Business models, decision processes and economic analyses of dynamics in logistics.
- Holistic interdisciplinary methods for modeling, analysis, and simulation of dynamics in logistics.
- Adaptive and dynamic control methods in logistics.
- Synchronization of material, information, decision, and financial flows.

The implemented supporting measures are courses, training, and coaching, most of them related to improving so-called soft skills, e.g., to gain more awareness on the impact of cultural differences and languages in scientific as well as social cooperation (Zhang et al. 2015). The IGS is combining scientific content with personnel and personality development in all offered measures as a “hidden agenda” (Rügge and Klempien-Hinrichs 2013).

In 15 years, 100 courses or organizational course units were conducted, in which nearly 300 scientists participated; that equals 1550 participations in total. The majority of the participants are doctoral candidates, i.e., scholarship holders of the IGS or research assistants of Log*Dynamics*. However, international guests from several mobility programs, not all of them conducting a PhD, have also participated in the courses of the IGS.

At this current juncture, 76 young researchers from twenty-four countries started their doctorate with a scholarship at the University of Bremen within the framework provided by the IGS. Fifty-three of them have been awarded a doctoral degree

(IGS 2021). They are professionally supervised by LogD*ynamics* professors in one discipline and also submit their dissertation in that discipline. The majority of the doctoral theses have been submitted to the three Engineering faculties Production Engineering (21), Electrical Engineering/Physics (12), and Mathematics/Computer Sciences (8) of Log*Dynamics*; ten to Business Administration/Economics. As diverse and individual as doctoral topics are, they have always had a strong relation to logistical problems. Among others, a significant scientific contribution was made to autonomous logistics as well as to the modeling and optimization of different kinds of supply chains, to the modeling of uncertainties and trust, to AI-based decision support, to protocols of mobile communication for Industry 4.0 applications, to sensors and sensor networks for food transports.

Through the measures of the IGS, however, there is a continuous moderated interdisciplinarity during the average 4-year doctoral phase. In addition, the provided customized continuing education courses are accessible to all Log*Dynamics* scientists, regardless of their contracts. Researchers pursuing their doctorates within a traditional assistance doctorate as well as postdocs, professors, and guest researchers are addressed by the offered courses. The orientation and composition of the courses is not left to chance but is also guided with regard to an interdisciplinary and international mix.

Sixty-seven of the 100 courses offered by the IGS have been evaluated by the participants. Quantitative evaluations as well as qualitative feedback were requested using a standardized questionnaire. The aim of each evaluation is to improve the IGS offering, so there were also queries. The evaluation was therefore not anonymous. This had as a positive effect an above-average return of answers as well as detailed answers and discussions of the organizer on HOW to adapt the course to the current needs with participants and lecturers. On average, between 11 and 23 people participated in a course in 12 years under consideration. The proportion of IGS doctoral scholarship holders varied from 24% to 56% and averaged 37.5% overall. The proportion of women in this period varied between 31% and 54%, for an average of 43%. On a scale of 1 (excellent) to 6 (poor), 88% of the evaluated courses received an average rating better than 2.5, meaning that only eight courses received a lower rating. 41 were rated in the upper range of 1.

All offers of the IGS are not general, they always undergo a specific adaptation that takes into account the realities of logistics, i.e., the related disciplines, the international cooperation, and the composition of the participants. For this purpose, the lecturers are also sensitized to these specifics in advance and during the implementation of the courses, and the specifics are made explicit. The qualitative study introduced in the second part of this paper emerged from this sensitization.

3 Part 2: Study on International, Interdisciplinary Education

International education, in its many forms and fashions, provides meaningful experiences and leaves lifelong impressions to students who engage in it. Concurrently, educators who are engaged in this endeavor also find much meaning in developing programs and interacting with the students. This cooperative paper explores the experiences of two educators involved in a doctoral summer program on professional writing as a component of an international education experience of engineers within a partnership between the United States, Ethiopia, and Germany. Unique to this partnership were (a) the interdisciplinary nature of the endeavor, (b) doctoral candidates from different disciplines from Ethiopia, (c) professors from the United States, (d) an international, interdisciplinary doctoral training group housed at a university in Germany, and (e) the summer course that coalesced the international experience. This paper will explore the perceptions of the professor from the College of Education at Texas Tech University (TTU) in the United States and the managing director of the International Graduate School for Dynamics in Logistics (IGS) in Germany.

3.1 International Educational Experiences

Gelpi (1985) espouses intercultural activities in higher education inevitably furnishes students and teachers with greater levels of confidence. He posits that intercultural activities involving higher education institutions that focus on global collaboration are beneficial to the students, the institutions, and the environments they influence. With these similar notions as Gelpi as the undergirding notions, the College of Engineering and the College of Education at TTU embarked upon collaborating with Jimma University in Ethiopia to develop graduate degrees in engineering and teach many of the courses therein. The impetus was not only to create international collaborations but to enhance the educational landscape of Ethiopia by the creation of science teachers and university professors through graduates of the program. One significant component of the program was to provide the Ethiopian students with an international experience outside of their home country. It is through this component that the IGS and the BIBA became an important and indispensable element. Between July 2017 and July 2018, 18 Ethiopian doctoral students traveled to Bremen, Germany, and engaged in academic international experience for three months at the “TTU PhD SummerCamp.” BIBA served as administrative host of the three months of international experiences of Ethiopian doctoral candidates and their American professors; the IGS managed all organizational issues and developed the program in Germany.

Another crucial component of the program was to develop the students’ skills in professional and publication writing. These two components gave rise to the interdisciplinary writing intensive course within the TTU PhD SummerCamp. This “Research Seminar” with the title “Proposal Writing: Professional Presentations and Publications” was implemented for two consecutive years. The faculty member from the College of Education at TTU designed it; and the experiences and observations of the facilitators of the program encapsulated and exemplified intercultural cooperation among different countries. Some parts of the curriculum of the course have been by email before and after the presence phase. In the first year, the scientific managing director of the IGS joined some classes of the research seminar. During and after the course, both educators discussed their experiences and started comparing the American and German system of supporting doctoral candidates. That was the moment to start the research and cooperation for this chapter. In the second year, there were less Ethiopian doctoral candidates left for an international experience as well as for participating in the Research Seminar. Therefore, the educators considered opening the course to doctoral candidates from Log*Dynamics*. An adaption of the Research Seminar was necessary to meet both needs: the needs of the Ethiopian doctoral candidates of the American system and the needs of those who are going to finish their doctorates in the German system.

3.2 TTU PhD SummerCamp Research Seminar

The Research Seminar course was developed to engage students in activities related to writing research proposals, presenting their research, and preparing their research for publication at the doctoral level, which included dissertation proposal preparation. Although the development of specific presentation and writing skills stood at the core of the course, the faculty member of TTU lectured on issues surrounding proposal structure, meaning, and purpose of the proposal, understanding a professional and scientific audience, preparing research for publication, and preparing research for professional presentations. Imbedded in the course were activities for the students to critically think about their role as a researcher and develop strategies to enhance their presentation, mentoring of others, and pedagogy. Over the 2-year span, the course was amended from a one-week course to a 2-week course. The first year only Ethiopian students were enrolled; and the second-year doctoral candidates from both Jimma University and the University of Bremen were enrolled. The observations by the faculty developer and the IGS managing director serve as the primary data source for this piece.

4 Qualitative Research Methods

Through scholarly personal narratives (SPN) (Nash 2004), the perceptions of the two university professionals associated with a German-USA educational partnership were explored relating to their experiences with the development and execution of an academic program for Ethiopian and Bremen doctoral students. Gorichanaz (2017) posited, "the collaborative, phenomenological approach to self-study [can] to allow for deep personal reflection while also stimulating open conversation among researchers and educators" (p. 4). An integral aspect of the design was plural positionality (Louis et al. 2017) since the authors are the above-mentioned university professionals and served as both (a) the participants, key informants, who articulated their experiences, and (b) researchers, who analyzed the meaning of their experiences. The data from the written narratives served as the primary data. Through the data analysis process, meanings of the experiences are developed and explicated (Nash 2004; Patton 2014; Lincoln and Guba 1985).

The research questions (RQs) that guided this narrative inquiry were: What are the perceptions of the two interacting educators about:

- RQ1: the inception of the international partnership and the program?
- RQ2: benefits and challenges of the international partnership and the program?
- RQ3: the vision and aspiration of the international partnership and the program?
- RQ4: the experiences of students who participated in the international partnership and the program?
- RQ5: the multidimensional features of the international partnership and the program?

4.1 Procedure and Analysis

An instrument consisting of six items/prompts was developed (Appendix A) for the researchers, who were the university professionals, to reflect and compose their narratives on their perceptions two university professionals associated with a German-USA educational partnership. The prompts were framed from the RQs focused on their experiences and observations as educators involved in a multi-disciplinary/multidimensional international collaboration between two universities. Before the instrument was distributed to the university professionals, it was vetted by the primary researcher to an independent qualitative researcher to make certain that the prompts were accurate and clearly addressed the fundamental nature of the study. The university professionals were given four weeks to reflect on the prompts, make notes and write sections, edit and re-edit, and construct their narrative with no limits on expression or articulation.

The narratives were distributed to each researcher, who were the university professionals involved in the TTU PhD SummerCamp, for individualized independent analysis for open coding. Subsequently, they met via Zoom to discuss their findings, disaggregate the data, delve further into the meanings of the narratives, and clarify the expressions. It provided opportunities for member-checking whereby clarity of meaning and ensuring the accuracy of the information was the central activity. The process of axial coding in the data analysis linked the common connections between the participants' responses showing shared observations and mutual perceptions about the events, individuals, experiences within the program (Saldaña 2015). From this exercise, themes emerged from the narratives. Additionally, the study ascribed to Lincoln and Guba's (1985) concept of trustworthiness. The self-analysis design and analysis discussions provided opportunities to further explain and clarify their meanings and experiences. This member-checking and reevaluation enabled the researchers to extricate very detailed and pertinent meanings of experiences with relation to their experiences and perceptions.

4.2 Positionality

Positionality is vital when communicating individuals' life experiences, insights, observations, and expressions of realities (Bourke 2014). Holmes (2020) stated, "The term positionality both describes an individual's world view and the position they adopt about a research task and its social and political context" (p. 1). Hence, being cognizant of the researcher's context is key to understanding their lens when engaging in the process, data, analysis, and theme development. And for SPN their roles as key informants call specifically for a comprehension of their backgrounds.

Dave is a tenured associate professor in higher education at a large public research university in the United States. He has over twenty years of experience in American higher education, holding positions ranging from director of student success center to executive director of a large academic program to tenure-track faculty. He has spent 8 years as a faculty member in the College of Education at Texas Tech University and is currently embarking on his initial term as a faculty member at the University of Houston. He is originally from Trinidad in the Caribbean but has lived the majority of his life in the United States. He identifies as an Afro-Indo-Trinidadian.

Ingrid is a managing director at a young, medium-sized public research university in Germany. She worked in different positions: Firstly, eight years in technology transfer at an interdisciplinary Research Center at the University of Bremen. For 14 years, she has been working as coordinator of an international doctoral training program of an interdisciplinary research cluster. She guided more than 50 doctoral candidates of logistics from 24 different countries successfully through their doctorate procedures in five different disciplines at the University of Bremen.

5 Findings

By now, two major themes emerged from the analysis of the narratives: (1) Bunter Blumenstrauß and (2) Academic Approach Convergence. The themes directly address the two university professionals' perspectives with a German-USA educational partnership. It reflected their insights, experiences, observations, interactions, and involvement with the TTU PhD SummerCamp. Although qualitative work is not generalizable, the findings can indeed shed light, provide perspective, and create greater awareness when developing cross-national and international educational programs.

5.1 Theme 1: Bunter Blumenstrauß

"Bunter Blumenstrauß" roughly translates into a "colorful bunch of flowers." Throughout the narratives, the wondrousness of the multiplicity of people, mixture of cultures, and differing educational systems converging in one place were expressed. The TTU PhD SummerCamp possessed one academic goal, but concurrently embraced and utilized the differences within the cultures and systems to benefit the students' learning outcomes. Thus, it was a symphony and cacophony of cultures wrapped under the banner of education. The beauty of the program, as well as some of the struggles within the program, was rooted in the diversity of the German, American, Ethiopian, and other cultures engaged in a common academic journey.

Dave articulated that the multicultural components of the program were what enriched the program and was one of the central aspects of the design. He stated,

> part of the core aspiration was the exchange of ideas and culture internationally. The program is one meant to expose all the students and faculty involved to different cultures, ways of thinking, and different educational systems with the hope that the African students will ultimately benefit from both the academic program and cultural exposure.

Within the design, Ingrid indicated that the Ethiopian students in particular had to adjust to German culture because they immersed for an extended period of time. Although the program was a magnificent opportunity for the students. She stated,

> Most of the Ethiopian doctoral candidates had never been abroad nor in a Western country before the TTU PhD SummerCamp in Germany. Their stay was limited to three months; thus, they had to adapt to daily life very quickly. The project brought three cultures (and languages) together...

However, what is important to note was her belief that three months may not have been sufficient for students to adapt to the culture of both the host country and the academic and social expectations. Ingrid indicated,

> Gaining international experience takes time! Three months is not sufficient to break someone out of her/his comfort zone, particularly not when [individuals start] in a homogeneous [cultural] group. It takes much longer to generate (automatic) awareness for intercultural, international, interdisciplinary and diverse encounters.

Dave, however, did have an instance in which he engaged with an Ethiopian student outside of the classroom and observed the importance of the international collaboration. Dave shared,

> I believe that the students developed through both the academic courses and the cultural exchange. I rode the train one evening with one of the students and the conversation we had about what he learned from being Germany was amazing. Plus, he seemed comfortable navigating life in Bremen. He was comfortable on the train and in the city and expressed his enjoyment in living in an apartment in the city.

Through both narratives, it was clear that the TTU PhD SummerCamp offered a great benefit to the students involved. However, within any bunch of flowers thorns/Dornen exist. With the amalgam of cultures, there were instances when there were some complications with respect to academic, societal, and cultural expectations. Ingrid shared,

> [It] was a challenge for the Ethiopians even if they received tremendous support from the hosting institution. They stuck together most of the time and didn't make any friends with

> Germans, neither the first group nor the three people of the second year. Integration was impossible.

There were instances of language barriers, academic expectation parries, and societal expectations barriers. Dave explained,

> The two biggest challenges were language and academic expectations. Most of the class instruction was conducted in English. The instructor's native tongue was English (British) and operates within an educational system that is also English (USA). The two main coordinators of the program both spoke languages other than English (German and Kenyan), the Ethiopian students spoke their native Ethiopian language, the German students spoke German. Thus, there were times when explanation of 'technical terms' was necessary and unexpected.

Although in the first year, the primary languages were rooted in the German, American, and Ethiopian cultures, in the second year with the inclusion of the Bremen doctoral students, there were people who were Chinese and Indonesian. So, beyond the confluence of German, American, and Ethiopian academic structures was the additional layer of multiple native languages of the students and the exposure to multiple academic higher education systems. Overall, the Research Seminar provided a platform and landscape for multiple cultures to be engaged in an academic program for doctoral students and include three countries in the development. Although the cultural exchange was extremely beneficial and was one of the most admirable aspects of the TTU PhD SummerCamp, it did have elements that resulted in some difficulty.

5.2 Theme 2: Academic Approach Convergence

It was clear that multiple academic systems would be converging during the development and operationalization of the TTU PhD SummerCamp. The curriculum and degree outline for the Ethiopian doctoral students at Jimma University were American and crafted as such. The students matriculated undergraduate and master's degrees in Ethiopia under the auspices and rubric outlined by the Ministry of Education in Ethiopia. However, for organizational reasons, the TTU PhD SummerCamp took place at the University of Bremen in Germany. In the second year, with the addition of the students pursuing doctoral degrees in a German system, the differences in academic approaches became more apparent. Thus, there was an inherent convergence of academic paradigms that was expected. But this challenge also allowed both university professionals to examine and learn about other educational structures.

Ingrid aptly expressed that both she as a professional and the students would be able to garner a glimpse into other educational systems. She shared,

> There was the chance for a few doctoral candidates of the German system to benefit from a regular course of the US system for doctoral candidates… [and] the managing director of the IGS had the opportunity to look into the US education system for doctoral candidates.

Concurrently, Dave was able to realize through both cycles of the program that,

> The structure of the thesis or dissertation and academic publishing was foreign or strange to many of the participants. Understanding the rationale and form of writing was complicated. Many students viewed it as discipline specific differences but as I examine it further it was more an educational system expectation that may not be common in Ethiopian or German systems. Thus, the differences within the system were highlighted in the academic writing.

Additional to the differing academic structures was the overlap of closely related yet independent disciplines represented in the classroom. Ingrid noticed,

> Multi-disciplinarity was not intended specifically in the program. It was a coincidence that already the first, culturally homogeneous group of participants of the Research Seminar brought doctoral topics from different disciplines. As a result, each of them also had a different scientific background. For the second course, it was clear that there would be disciplinary diversity as well as participants from several different nationalities.

Thus, emanating from the analysis of the observations was the idea that although international partnerships may have overt comprehension of cultural confluence, they could be more intentional at the outset to explore the academic approach differences between educational systems and models.

Both Dave and Ingrid observed and commented on the differing research paradigms particularly between the American and German systems. The German doctoral approach in Engineering was one that starts the research path mostly with "field, project and experience"-based questions of the real world, whereas the American model seemed steeped more in the theory surrounding research and literature reviews before thoroughly noticing the real world's evidence and starting engaging in fieldwork. Both professionals agreed that the expected and delivered outcomes (theses) were the same but the approaches in educating students are different. And this difference became evident in the writing process, which was at the central focus of the Research Seminar. Ingrid stated,

> In Germany, there is no tradition to train how to write down one's research idea. In the German research tradition in engineering, writing about one's research holds second place behind 'carrying out' the research proper. In the US system (and therefore valid for the Ethiopian doctoral candidates of this Research Seminar) it seems that students first need to write their ideas down in a structured form and then the ideas will be evaluated whether they are worth being followed through. Basically, both groups need the same course contents about writing. The mindset in which these contents are approached makes it easier or harder on students to profit from the course. In the second year group, the Ethiopians absorbed more from Dave than the student from Bremen. The latter reflected Dave's teachings and ended up somewhat opposed since the Research Seminar seemed to be 'just about pretty words'.

Dave explained his observations,

> The structure of the thesis or dissertation and academic publishing was foreign or strange to many of the participants, especially those from the German system. Understanding the rationale and form of writing was complicated. Many students viewed it as discipline specific differences but as I examine it further it was more an educational system expectation that may not be common in Ethiopian or German systems. Thus, the differences within the system were highlighted in the academic writing... One way I addressed this was through my pedagogy. My delivery always engaged the students in manners in which specific

> disciplines would express their research. Students had to find published articles to also understand the writing style and flow of their area. Overall, it was not difficult to merge disciplines. My experience is to teach writing across disciplines, so it was not difficult for me, but students sometimes had difficulty making sense of certain aspects of the writing. And some did not see the connection between their work and having to express it to an unknown audience.

The Research Seminar for the TTU PhD SummerCamp in numerous ways challenged the thinking and pushed the paradigmatic barriers of both the students and educators. It caused some level of dissonance where individuals had to wrangle and discover the academic space between "possessing knowledge" and "expression of knowledge." Students had to step outside of their comfort zones when addressing how to express their research, and the university professionals had to find ways to foster meaning-making between "knowledge" and "academic articulation." This becomes even more complex when culture, language, and multidisciplinarity are part of the classroom ecology.

6 Conclusion

According to Quacquarelli Symonds [QS] (2019), international partnerships "have contributed endlessly to academic and scientific progress" (para. 10) and benefit the students and the institutions on several levels. Furthermore, these partnerships cultivate research opportunities, cultural awareness, international experiences, curriculum development, and degrees formed in collaboration with partner institutions (QS, 2019). The partnership between TTU and the University of Bremen with the TTU PhD SummerCamp provided all of those elements to Ethiopian doctoral students from Jimma University and German, Chinese, and Indonesian doctoral researchers at the University of Bremen. This narrative study outlines two components that the program provided and sheds light on the complex nature of such a partnership. The program provided both the students and the educators with the opportunity to expand their horizons both culturally and academically. The IGS organized living accommodation, cultural and professional tours, and social opportunities for the Ethiopian students to interact with members of the German university and engineering community. Differences in cultural expectations and contrasting cultural mores, norms, and folkways were some of the obstacles that individuals in the program observed. Language barriers, contrasting academic structures, and multidisciplinarity posed challenges particularly in terms of comprehension of what was needed to be fully successful in the Research Seminar. However, despite these drawbacks, the researchers, who were involved in both coordinating the program and teaching the course, assert that the program overall was beneficial to the students and the German and American universities. The authors posit that engaging in long-term cultural discussion about academic expectations should be a key component in the development and strategic planning of any international partnership. Fostering a greater understanding of the academic structures of the countries' university systems, the differences in academic approaches with respect to attaining degrees, and delving into the philosophical underpinnings of the partner universities' curriculum are crucial to consider in program development. Additionally, cultural competency should be a component of both students' and professionals' preparation for the program. For the TTU PhD SummerCamp, the professor for the United States spoke English fluently as his native language and possessed a marginal grasp of German. The German coordinator spoke German fluently as her native language and possessed a good grasp of English. Neither spoke any Ethiopian language. The command of English of the Ethiopian students was medium, none of them spoke any German language. Not feeling comfortable with a (foreign) language may be a barrier. The unexpressed assumptions and implicit rules hidden in the foreign system cause a much higher barrier. Therefore, universities who are considering developing international partnerships should seriously consider providing educational professional development of their faculty and staff with respect to getting aware of the culture of their partner institutions. Furthermore, educators of international students (on all levels of education) should receive the opportunity to train their perception of culturally founded differences and to make it explicit. Training and experiences are needed to develop these abilities before one can share or teach them.

7 Outlook

This example of the German-USA international partnership between TTU and the University of Bremen exemplifies the importance of intentional program design, the crucial nature of identifying the distinctions between university systems, and the acknowledgment of the continued benefit of opportunities for cultural exchange across, between, and amongst nations of the world. The cooperation on educational research of both of the educators involved in the study will be extended by a Fulbright-funded research stay of Dave Louis at the IGS in 2022. Furthermore, the IGS is sharing its experiences with the Asian partners of the Erasmus+ Capacity Building project SSAPI as well as with the international partners gained by the alumni of the IGS.

The interdisciplinary research focus of the next incoming doctoral candidates of the IGS will be closely related to those of Log*Dynamics*. The trend toward individualization has taken hold of logistics. The rapid growth of e-commerce in the current Corona pandemic has given this trend an additional boost. Increased mobility and more packaging material are the results. On the other hand, the effects of climate change are now evident, requiring logistics to reduce environmental toxins and improve the robustness of logistics supply chains. This raises new research questions whose treatment is of international interest and can only be answered meaningfully in interdisciplinary cooperation. This applies in a similar way to production, civil, or construction engineering. Therefore, the examination of factors of successful international interdisciplinary education and identifying best practices will be continued and intensified.

Acknowledgments The TTU PhD SummerCamp was part of the TTU project "Postgraduate Programs for Civil Engineering and Construction Technology." It has been funded by The Ministry of Education of the Federal Democratic Republic of Ethiopia (through the British Council and with the financial support of the German Development Cooperation (GIZ)). Many thanks go to Dr. Renate Klempien-Hinrichs who helped Ingrid with several methods to put her thoughts into a precise structure and English sentences.

A.1 Narrative Prompts

A.1.1 Narrative Prompt 1: Inception of the Partnership and the Program

As you think about the creation of the international partnership what were the main motivating factors in your mind for the program to be created? Why was it important for this partnership and program to be created? What was the core purpose of the program to be established? Who was it supposed to benefit? After it was completed (both year 1 and year 2) were the goals of the program achieved? Why would you say they were or were not achieved? What were the most positive aspects of working with interactional partners? What were the most negative aspects of working with the international partners?

A.1.2 Narrative Prompt 2: Program Vision and Aspiration

What did you envision this program to be for your university/unit? What did you aspire this program to be for the students involved? What were the outcomes you had hoped for during its creation (for the students and the administrators involved)? What did the partners express they wanted from the program? Was what the administrators wanted different from what you wanted as an educator?

A.1.3 Narrative Prompt 3: Challenges

Were there challenges in the creation of the program? Describe any challenges you had working with a university from another country? How were these challenges overcome? Describe any challenges you had with students?

A.1.4 Narrative Prompt 4: Student Experiences

How would you describe the overall experience of the students who participated in the program? Describe what you believe were some of the most positive experiences that the students had while participating in the program? Describe what you believe were some of the most positive benefits that the students acquired while participating in the program? Describe what you believe were some of the most negative experiences that the students had while participating in the program?

A.1.5 Narrative Prompt 5: Multidimensional

How did the idea of multidimensional/multidisciplinary education influence the development of the program? Were there challenges with the merging of the four disciplines? How did you overcome or address the challenges associated with being multidimensional? What did you have to explain or translate for your international partner about the differences in educational expectations between the two countries? Were there language barriers about educational terms? Describe these barriers? Share how the barriers were overcome?

A.1.6 Narrative Prompt 6: Conclusion

Now that the program is completed, where are the most beneficial aspects of the program? What was the most positive aspect of the program? What elements of the program would you like to see replicated in other programs? What have you learned, as an administrator, from being involved in this program? What have you learned, as an educator, from being involved in this program?

References

Bourke, B.: Positionality: Reflecting on the research process. Qual. Rep. **19**(18), 1–9 (2014)

Gelpi, E.: Lifelong Education and International Relations. Croom Helm, London (1985)

Gorichanaz, T.: Auto-hermeneutics: A phenomenological approach to information experience. Libr. Inf. Sci. Res. **39**(1), 1–7 (2017). https://doi.org/10.1016/j.lisr.2017.01.001

Holmes, A.G.D.: Researcher positionality – A consideration of its influence and place in qualitative research – A new researcher guide. Int. J. Educ. **8**(4), 1–10 (2020). https://doi.org/10.34293/education.v8i4.3232

IGS: Alumni of the IGS (2021). Retrieved from http://www.logistics-gs.uni-bremen.de/alumni.html?&L=1 (last visit July 25, 2021)

Lincoln, Y.S., Guba, E.: Naturalistic inquiry. Sage, Newbury Park, CA (1985)

Louis, D.A., Thompson, K.V., Smith, P., Williams, H.M.A., Watson, J.: Afro-Caribbean immigrant faculty experiences in the American Academy: Voices of an invisible black population. Urban Rev. **49**(4), 668–691 (2017). https://doi.org/10.1007/s11256-017-0414-0

Nash, R.: Liberating Scholarly Writing: The Power of Personal Narrative. New York (2004)

Patton, M.Q.: Qualitative Research & Evaluation Methods. Sage, Thousand Oaks, CA (2014)

QS [Quacquarelli Symonds]. Why are International Collaborations so Important For Universities? (2019). Retrieved from https://www.qs.com/why-are-international-collaborations-so-important-for-universities/

Rügge, I.; Himstedt, A.: Multidimensionale Diversität in der Logistikausbildung. In: Industrie 4.0 Management, 2/2015, GITO Verlag, Berlin, pp. 61–65 (2015)

Rügge, I.; Klempien-Hinrichs, R.: Coping with Multi-Dimensional Diversity in Logistics Education in Proceedings of 6th IFAC International Conference on Management and Control of Production and Logistics (MCPL), Fortaleza, Brazil, September 11–13, 2013, pages 456–462, (2013) IFAC-PapersOnline, DOI: https://doi.org/10.3182/20130911-3-BR-3021.00101

Rügge, I., Scholz-Reiter, B.: Internationalization of the Doctoral Training Program for the German Degree “Dr.-Ing.”, In: Proceedings of SKIMA 2011, Benevento, Italy, IEEE (2011). https://doi.org/10.1109/SKIMA.2011.6089998

Saldaña, J.: The Coding Manual for Qualitative Researchers. Sage, Thousand Oaks, CA (2015)

Scholz-Reiter, B.; Rügge, I.: Promovieren im Forschungsverbund LogDynamics: international + interdisziplinär = vernetzt. In: Hippler, Horst (Hrsg.): Ingenieurpromotion – Stärken und Qualitätssicherung. Heidelberg u.a.: Springer, S.115–124 (2011)

TTU SummerCamp. TTU PhD SummerCamp in Bremen (2018). Retrieved from http://www.logistics-gs.uni-bremen.de/467.html?&L=1 (last visit July 25, 2021)

Universität Bremen: Mission Statement (2021). Retrieved from. https://www.uni-bremen.de/en/university/profile/mission-statement (last visit July 25, 2021)

Zhang, H., El-Berishy, N., Zastrau, D., Marwat, S.N.K., Tan, Y., Palafox, J., Rügge, I.: Interdisciplinary perspectives on knowledge management in logistics. In: Dethloff, J., Haasis, H.-D., Kopfer, H., Kotzab, H., Schönberger, J. (Eds.): Logistics Management: Products, Actors, Technology – Proceedings of the German Academic Association for Business Research, Bremen, 2013 (Lecture Notes in Logistics). Springer, pp 491–504 (2015)