Lecture Notes in Logistics

Series Editors: Uwe Clausen · Michael ten Hompel · Robert de Souza

Christian Bierwirth
Thomas Kirschstein
Dirk Sackmann *Editors*

Logistics Management

Strategies and Instruments for digitalizing and decarbonizing supply chains - Proceedings of the German Academic Association for Business Research, Halle, 2019

Springer

Lecture Notes in Logistics

More information about this series at http://www.springer.com/series/11220

Christian Bierwirth · Thomas Kirschstein ·
Dirk Sackmann
Editors

Logistics Management

Strategies and Instruments for digitalizing and decarbonizing supply chains - Proceedings of the German Academic Association for Business Research, Halle, 2019

Springer

Editors
Christian Bierwirth
Department of Business Administration
Martin Luther University Halle-Wittenberg
Halle (Saale), Sachsen-Anhalt, Germany

Thomas Kirschstein
Department of Business Administration
Martin Luther University Halle-Wittenberg
Halle (Saale), Sachsen-Anhalt, Germany

Dirk Sackmann
Department of Business Administration
and Economics
University of Applied Sciences Merseburg
Merseburg, Germany

ISSN 2194-8917 ISSN 2194-8925 (electronic)
Lecture Notes in Logistics
ISBN 978-3-030-29820-3 ISBN 978-3-030-29821-0 (eBook)
https://doi.org/10.1007/978-3-030-29821-0

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 contains selected papers presented at the 11th *Logistics Management* conference (LM 2019) of the Scientific Commission for Logistics (WK-LOG) of the German Academic Association for Business Research (VHB). The LM conference series is continued every two years at different places in Germany. It aims at providing a forum for scientists and practitioners in business administration, IT, and industrial engineering to present and discuss new ideas and technical developments related to the management of logistic systems. LM 2019 was hosted by the Martin-Luther-Universität Halle-Wittenberg in cooperation with Merseburg University of Applied Sciences. It took place from September 18–20, 2019, in Halle (Saale). Previous LM conferences were held in Bremen (1999, 2013), Aachen (2001), Braunschweig (2003, 2015), Dresden (2005), Regensburg (2007), Hamburg (2009), Bamberg (2011), and Stuttgart (2017).

The LM 2019 conference put a special focus on the digitalization of supply chains and recent attempts to decarbonize the transport industry. Not only do both fields address vitally important concerns bringing out (and together) many new ideas, innovative concepts, and future visions; the implementation of digitalized supply chains and the application of electrified mobility are also happening very quickly. Considering a snapshot in 2019, the variety of possibilities and ways toward digital supply chains and green transport processes appears amazing. To give an insight into the field, LM 2019 has invited three keynote speakers to examine ongoing developments:

- Ola Jabali (Politecnico di Milano): Emission-oriented modeling and optimization in transportation,
- Michael Knemeyer (Ohio State University): The importance of relationships in the age of digitalization,
- Jan Fabian Ehmke (Otto von Guericke University Magdeburg): Big data for transportation optimization.

In addition to the keynote talks, around 50 presentations were given at LM 2019 of which 21 are printed as full papers in this proceedings. These papers have been selected from 25 full paper submissions in a careful review process involving three referees for each paper in up to two rounds of revision. The accepted full papers address a broad spectrum of facets of logistic systems with regard to digitalization, sustainability, and optimization. They divide this book into three parts, considering the digitalization of supply chains, sustainable supply chain management, and supply chain operations. We hope it provides insights into the state of the art of logistics management and, thus, stimulates future research.

September 2019

Christian Bierwirth
Thomas Kirschstein
Dirk Sackmann

Acknowledgments

The editors express their gratitude to all of the authors and to everybody who has contributed to this volume. In particular, we thank Springer for the easy and uncomplicated collaboration in the editing and publishing process. Moreover, we thank the German Research Foundation (DFG) for granting traveling funds which allowed us to invite internationally renowned keynote speakers. Our special gratitude goes to DHL for awarding the best full paper contribution of LM 2019, and to InfraLeuna GmbH, Arvato Bertelsmann distribution GmbH, and the city of Halle (Saale) for their generous support of LM 2019. Furthermore, we gratefully acknowledge the efforts of the program committee for reviewing the contributions submitted to the conference. The LM 2019 program committee consists of

Prof. Dr. Christian Bierwirth, Martin-Luther-Universität Halle-Wittenberg
Prof. Dr. Ronald Bogaschewsky, Julius-Maximilians-Universität Würzburg
Prof. Dr. Udo Buscher, Technische Universität Dresden
Prof. Dr. Jan Dethloff, Hochschule Bremen
Prof. Dr. Jan Fabian Ehmke, Otto von Guericke Universität Magdeburg
Prof. Dr. Kathrin Fischer, Technische Universität Hamburg-Harburg
Prof. Dr. Hans-Dietrich Haasis, Universität Bremen
Prof. Dr. Gudrun P. Kiesmüller, Otto von Guericke Universität Magdeburg
Dr. Thomas Kirschstein, Martin-Luther-Universität Halle-Wittenberg
Prof. Dr. Natalia Kliewer, Freie Universität Berlin
Prof. Dr. Herbert Kopfer, Universität Bremen
Prof. Dr. Herbert Kotzab, Universität Bremen
Prof. Dr. Anne Lange, Université du Luxembourg
Prof. Dr. Rudolf Large, Universität Stuttgart
Prof. Dr. Rainer Lasch, Technische Universität Dresden
Prof. Dr. Dirk C. Mattfeld, Technische Universität Braunschweig
Prof. Dr. Frank Meisel, Christian-Albrechts-Universität zu Kiel
Prof. Dr. Taïeb Mellouli, Martin-Luther-Universität Halle-Wittenberg
Prof. Dr. Dirk Sackmann, Hochschule Merseburg
Prof. Dr. Katja Schimmelpfeng, Universität Hohenheim

Prof. Dr. Jörn Schönberger, Technische Universität Dresden
Prof. Dr. Stefan Seuring, Universität Kassel
Prof. Dr. Thomas S. Spengler, Technische Universität Braunschweig
Prof. Dr. Axel Tuma, Universität Augsburg
Prof. Dr. Guido Voigt, Universität Hamburg
Prof. Dr. Carl Marcus Wallenburg, WHU Otto Beisheim School of Management

Contents

Digitalization of Supply Chains

Digital(ization) – A Single Construct Amidst Supply Management?

Cornelia Elsäßer, Andreas H. Glas(✉), and Michael Eßig

Bundeswehr University Munich, W.Heisenberg-Weg 39, 85577 Munich, Neubiberg, Germany
andreas.glas@unibw.de

Abstract. Digital technologies pervade products, services and companies. Keywords like Internet-of-Things, or big data analytics made their way on to almost every presentation and promotional slogan, seemingly becoming the backbone for every business organization, including supply management functions. However, the understanding of digital and related terms such as digitalization varies significantly. This research addresses the conceptual gap and seeks to identify characteristics and patterns of digital/digitalization. In addition, the terms digital, digitization and digitalization will be delineated. For this purpose, this research reviews existing literature. Digital as a phenomenon is structured through the application of deductive structure-discovering methods. The findings show the discrepancy between a technical and a managerial understanding. This paper proposes a conceptual model that structures the phenomenon into (1) digitization, (2) media convergence, (3) digitalization and (4) digital transformation.

1 Introduction – Antecedents of Digitalization

The increasing interest in digitalization is reflected by the growing amount of publications dealing with the phenomenon: entering digitalization as Google search term results in approximately 4 billion hits, of which 1.1 billion have been published in 2018, leaving the reader with a quite versatile understanding of digitalization (Google 2018). In addition, supply management literature discusses various aspects of digital, e.g. blockchain, web-based communication, or big data (e.g. Kache and Seuring 2017), however, experts still claim that "digital" still poses a major research gap in supply management, however claiming that digital technologies are becoming a game changer for current supply structures (Srai and Lorentz 2019). It is expected that digitalization will on the one hand have a major impact on how buyer-supplier interactions, on the other hand it might also reconfigure the value proposition of supply management, e.g. considering reciprocal dependencies and the (upstream) exchange of data.

In this regard, it becomes critical to distinguish how and in which ways digital affects current structures. Also, it remains unclear whether digital is to be treated as an adjective or noun, construct or concept, thus we will use digital as general term and interchangeably for the phenomenon prior to drawing a conclusion on the subject.

Digital per se is not a new phenomenon. As early as 1949, Forrester described the digital computer (Bennett 2002; Forrester 1949). Digitalization, in that sense, describes

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 3–18, 2019.
https://doi.org/10.1007/978-3-030-29821-0_1

the process of converting analogue signals into binary code. The present understanding of digital/digitalization goes far beyond that definition: besides the application of technologies and their increasing interconnectedness, changes in business processes and –models are nowadays also considered digital (Bracht and Masurat 2005; Hassani 2006; Henfridsson et al. 2009; Kreutzer and Land 2015; Pagani and Pardo 2017; Patnayakuni 2006; Yoo et al. 2012; Yoo et al. 2010a, b).

As rare as the considerations regarding the content factors of digital/digitalization are, as versatile are the effects reflected in research. While some authors provide a definition that is related to the original definition of Forrester, others highlight the connection of technical and social aspects (BMWi 2015; Capgemini Consulting 2018; Forrester 1949; Gartner 2018; KPMG 2018; Pagani and Pardo 2017; Tilson et al. 2010; Yoo et al. 2010a, b).

In order to illustrate the conceptual mess, Fig. 1 provides an overview of the range of definitions, based on content proximity of the different definitions in contrast to the original definition and scientific quality of the source. The scientific quality is based on the type of publication, with journal and conference publications ranked as highly scientific, monographs and working papers ranked as mediocre. Non-scientific- and practical research is ranked as having low scientific quality.

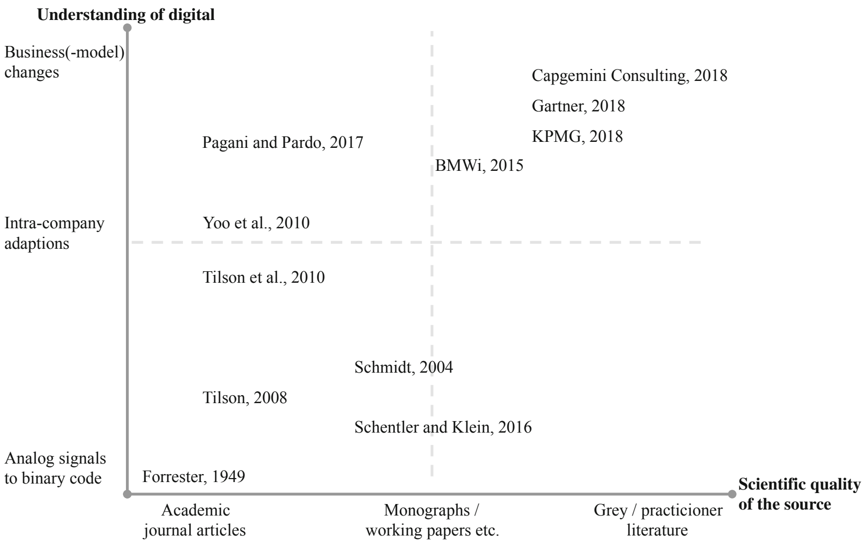


Fig. 1. Portfolio – proximity of definitions digital/digitalization/digitization.

Schentler and Klein or Schmidt describe digitalization as conversion of analogue signals to binary code (Schentler and Klein 2016, Schmidt 2004). The definition of Tilson contains the same elements as the one of Forrester, however changing the term to digitization indicates that digitalization has a different meaning than the original understanding (Tilson 2008). This becomes apparent when considering the contribution of Tilson et al. published at a later stage, where digitizing is described as a technical process, as opposed to digitalization, being described as sociotechnical process (Tilson et al. 2010). Yoo et al. come to the same conclusion when describing digitization as encoding of analogue information, while defining digitalization as process that “involves organizing new socio-technical structures” (Yoo et al. 2010a, b). The German Federal Ministry of Economics and Energy (BMWi) similarly describes digitalization as networking of all areas of business and society, as well as the ability to collect, analyze and implement relevant information, thus focusing on the socio-technical component of digitalization (BMWi 2015). The definition of Pagani and Pardo even goes beyond the combination of technical and social aspects, defining “digital transformation as the digitalization of previously analog machine and service operations, organizational tasks, and managerial processes […] in order to drive new value for customers and employees” (Pagani and Pardo 2017). This definition addresses foremost the recipients of digitalization efforts: customers and employees. Moreover, consulting firms such as Capgemini Consulting, Gartner or KPMG consider digitalization as key aspect for creating new business models based on the use of digital technologies (Capgemini Consulting 2018; Gartner 2018; KPMG 2018).

The classification of definitions as stated above indicates that there is no common understanding of what the content of digital/digitalization really is. Due to this gap, an operational definition or plan on how companies or functions can define and assess their status quo of digitalization is not yet available. Procurement and supply management, as the interface to the upstream market, is no exception to this shortfall, as electronic procurement and enterprise digital transformation are around since the 2000s, promising competitive advantage through process optimization (Zhu et al. 2006), thus guiding traditional industries towards digital processes. In addition to that, the shift in customer-value propositions requires especially the procurement function to adapt the supplier side accordingly.

To clarify the digital construct, this paper aims to answer the following research questions (RQ):

RQ: Which constituent characteristics define digital/digitalization for SCM construc(s)?

To answer that question, a structured literature review has been conducted.

Additionally, network analysis was used in order to identify structures of the phenomenon. As a result, an operationalized definition of digital/digitalization, as well as associated constituent characteristics, were derived.

The remainder of this paper is organized as follows. The next section describes the methodological approach of identifying and evaluating data in order to derive structures and characteristics of the phenomenon, successively providing an evaluation of the results. In the third section, the results are consolidated and merged. Section 4 provides a summary of observations and implications, including a conceptual maturity model and some concluding remarks.

2 Digital – A Phenomenon Everyone Knows – Undoubtedly?

2.1 Methodology

A systematic literature review is conducted to collect data as a basis for theory elaboration, i.e. construct specification, construct splitting, and structuring sequence relations of digital/digitalization (Fisher and Aguinis 2017). The literature review follows the methodology of a multistage, systematic review (Denyer and Tranfield 2009; Easterby-Smith et al. 2015; Tranfield et al. 2003), including citation and credential tracking (Wohlin 2004). The methodological approach ensures completeness and reproducibility of the results.

Relevant papers were identified in the Ebscohost, ScienceDirect and Wiley databases, as these comprise most relevant academic journal sources in the purchasing and supply management domain. The databases selected do not only cover nearly all highly ranked purchasing and supply management journals, but they also contain multidisciplinary journals related to the topic. As "digital" will lead to more cross-industrial supply relations, e.g. between IT-firms and manufacturers, it seems necessary to connect the knowledge base of different research streams more closely in order to derive a holistic picture. As supply management-oriented publications mostly resulted in a deep dive of specific digital technologies, we thus opened the field from there into multidisciplinary research. Thus, the review contains material that goes beyond the scope of purchasing and supply management.

The search was conducted without timely restrictions (search terms were Digital* AND/OR Digitized, Digitalised, Digitalization, Purchasing, Supply Manag*; SCM, Suppli*, Automati*, thus the hits result from either combination or individual notation of truncated words). As the amount of hits exceeded one million, the search was narrowed down to searching the search terms only in title, keywords and abstract. Further, journals that focus on digital as component of a certain technology or medicine, such as Journal of Optics Communications, Optics and Lasers in Engineering were eliminated. Search was checked with the results of a previously conducted literature review on a similar topic (Elsaesser 2018).

In total, 3129 journal titles were reviewed, leaving 756 relevant for keyword-, and 287 relevant for abstract review. The number was drastically reduced due to irrelevant hits (e.g. journals that contained digital photography). Finally, the full text of 169 articles was read, leaving 42 hits that are relevant.

The analysis was split into several steps: First, the publications in the analysis are structured and categorized (Mayring 1991, 2010; Miles and Huberman 1994). The descriptive analysis contains relevant features including temporal distribution, journals, industry focus, research methods, and theories used (Briner and Denyer 2012).

Second, the content of the publications is evaluated qualitatively in an iterative process (Mayring 1991, 2010; Miles and Huberman 1994; Pescher 2009; Pratt et al. 2005). For this purpose, the publications are analyzed according to frequent and common topics targeted, thus constituting individual categories, using abductive reasoning for category building (Corbin and Strauss 2007; Miles and Huberman 1994; Pratt et al. 2006). Based on the preliminary categories the publications are reread, categories are refined and data is coded. This is done iteratively until all excerpts match a category. Finally, the results are connected with axial coding in order to identify cross-references between individual categories. The categories then serve as basis for network analysis (Bastian et al. 2009; Corbin and Strauss 2007; Miles and Huberman 1994; Pratt et al. 2006; Spiggle 1994; Vasudevan et al. 2016).

Using network analysis corresponds with a call for extension of the range of methods in purchasing and supply management research (Knight et al. 2016). To perform the network analysis, the publications are ordered hierarchically originating from the citation or reference level, i.e. based on the originally identified publications; each reference is assigned a level. This results in different levels, which in total represent the sum of publications. Due to the hierarchical order, evaluations can be carried out on the entire network or at a respective sub-level. The arrangement of the respective nodes represents the proximity (closeness centrality) of the nodes to each other, the edges correspond the betweenness centrality. The thickness of the edges represents the frequency of the connection between two nodes (Bastian et al. 2009; Borgatti 2005; Knight et al. 2016; Newman 2001; Okamoto et al. 2008; Vasudevan et al. 2016; Zhang et al. 2009; Zhang et al. 2010). The results of the network analysis serve as a mechanism for identifying the relevance of categories for describing the phenomenon, thus categories that belong to the cluster digital and that are more central in the network are considered more relevant for the phenomenon (Borgatti 2005; Mayring 1991). As a result, an analytical framework for constitutive characteristics of digital, as well as an operational definition for digital/digitalization can be derived (Borgatti 2005; Newman 2001; Okamoto et al. 2008).

2.2 Descriptive Results

In total, 42 publications have been considered relevant for descriptive evaluation. Eleven publications are highly ranked journal publications (VHB 2018), while 31 publications are either conference proceedings or monographs. Grey literature has not been considered within the research. Most publications have been published in operations research (16 publications), information systems management (11 publications) or organization/human resource management (9 publications), only six publications can be assigned to logistics, purchasing or supply management.

The increasing relevance of digital/digitalization is also reflected by the distribution according to the year of publishing: 39 publications have been published after 2001, thereof 30 after 2010. Also, 39 papers focus on non-empirical research considering digital as a phenomenon on network level. Three contributions consider digital as dyadic topic with focus on empirical research. Contingency theory, resource-based view, innovation diffusion theory and dynamic capabilities are external grand theories mentioned when describing the phenomenon.

2.3 Qualitative Results

The plurality stated in the first section of this paper pursues when evaluating the 42 publications: while some authors focus rather on technical aspects of digital (Tilson et al. 2010; Yoo 2010; Yoo et al. 2010a, b), others focus on describing changes in companies from a more general perspective (Samulat 2017). The analysis reveals that often authors (subconsciously) divide the phenomenon into different stages/phases (Bullock and Batten 1985; Poole 2004). To clarify the different definitions and phases, Table 1 provides quotations for each phase referring to technology, company, and business model dimension. The four main phases as displayed in Table 1 have been derived directly from the authors as mentioned in the table. However, Schönbohm and Egle (2017) do not provide a business model adaption for phase two – this phase is inherited in phase one.

Yoo et al. describe the changes of digitalization as *four* "waves" while Samulat or Schönbohm and Egle describe *four* maturity stages (Yoo et al. 2010a, b; Samulat 2017; Schönbohm and Egle 2017). Samulat name the stages as "digital personae" (Samulat 2017). It becomes clear that authors often distinguish four different levels of digital, yet the quotations indicate that the descriptions contain overlapping or heterogeneous content. The understanding is neither consistent in describing the content of each level, nor in creating a holistic model encompassing all phases and previous findings.

2.4 Content Network Analysis

In addition to Table 1 which only gives an overview of three different authors, the next section describes the network analysis which contains the keywords of all 42 journal contributions, thus offering a broader view on the content of digital and going beyond existing phase definitions adding some depth to the constituent characteristics. The analysis is conducted with Gephi, an open source software for network analysis (Bastian et al. 2009). As described in the methodology section, raw data has been analyzed iteratively in order to identify frequent and common topics (Corbin and Strauss 2007; Miles and Huberman 1994; Pratt et al. 2006). Figure 2 illustrates the results of the content-network created with Yifan Hu multilevel layout algorithm, which combines multilevel and force directed algorithm basics to reduce large network complexity (Hu 2006).

Based on the network measures described in the methodology section, the construct digital can be divided into two different clusters, based on their centrality within the network. The thickness of the edges represents the closeness of the connections thus the thicker an edge the more related the nodes. The average edge length has been calculated with 1.715 units, the diameter of the network is 4 units, the diameter of cluster 1 is 1.843 units, with an average edge length of 1.34, whereas the diameter of cluster 2 is 2.385, with an average edge length of 1.621. The measures support the validity of forming two separate clusters. Moreover, the average local clustering coefficient, which measures the degree to which a node is embedded in its cluster, is 0.742 for cluster 1 and 0.529 for cluster 2. The average global clustering coefficient is 0.422, thus indicating that the nodes of cluster 1 have a higher transitivity than the ones in cluster 2, even though the global clustering coefficient and the measures confirm that the nodes within each cluster are linked more closely than the overall network. The colors of the edges are used as heat map, thus indicating that the four most frequent connections from digital are technical and managerial dimension, inter-organizational systems design, integration and organizational culture.

Thus, the analysis results in **two separate clusters**, containing 12 categories. Using a clustering algorithm, the categories are assigned as follows:

Table 1. Phases/Stages of digital.

Phase 1		
Technology	Yoo et al. (2010a, b)	*"The first wave of digitalization does not bring any fundamental changes in the tightly coupled layer of product architectures"*
	Tilson et al. (2010)	*"[…] the ubiquity of small and powerful digital computers, and (2) the ubiquity of connecting those computers […]"*
Company	Samulat (2017)	*"Digital conservatives hesitate to change"*
Business model	Schönbohm and Egle (2017)	*"At the beginning there is often an unstructured approach to digitalizing a business model"*
Phase 2		
Technology	Yoo et al. (2010a, b)	*"In the second wave of digitalization, we begin to see the separation of devices, networks, services, and contents that have been tightly coupled in the past"*
Company	Samulat (2017)	*"Digital starters: low digital intensity in applying new technologies"*
Phase 3		
Technology	Yoo et al. (2010a, b)	*"In the third wave of digitalization, we begin to see the emergence of novel products and services through the "mash-up" of different media across different product architectural boundaries"*
	Tilson et al. (2010)	*"Digitization was now moving beyond paving of the cow paths to the building of a new highway system"*
Company	Samulat (2017)	*"Digital fans: some digital initiatives, but no maximization of use for the company"*
Business model	Schönbohm and Egle (2017)	*"Digital Islands: Based on benchmarking small digital business models develop, but no synergies within companies"*
Phase 4		
Technology	Tilson et al. (2010)	*"Linked to new breeds of enterprise wide infrastructures (built around ERP/CRM systems and databases) radically new business models were invented e.g. Amazon, eBay, Google, and Netflix"*
Company	Samulat (2017)	*"Digital experts: have created a digital culture and investments"*
Business model	Schönbohm and Egle (2017)	*"Digitalization strategy: holistic consideration of digital business model"*

- Cluster 1 (red nodes) contains the categories digital, information, integration, IT, organizational culture, virtual, and knowledge management.
- Cluster 2 (blue nodes) contains the categories innovation, inter-organizational systems, inter-organizational systems design, supply management, and technical and managerial dimension.

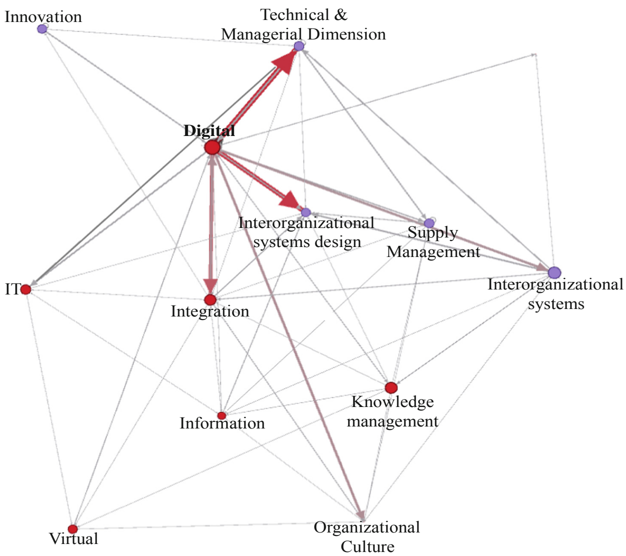


Fig. 2. Network analysis digital in purchasing.

The two different clusters illustrate that there is no profound consideration of digital in supply management literature, as the terms (supply management; digital) are located in different clusters. This depiction emphasizes that supply management talks about single synopsis of the phenomenon, but does not really target core aspects of it.

Moreover, when talking about digital, information and knowledge management play a central role, while a holistic view, considering technical *and* managerial aspects of the phenomenon, is still neglected. Categories located in cluster 1 have a higher proximity to the original (technical) definition of digital, while categories in cluster 2 tend to understand the phenomenon more broadly. On the other hand, the most frequent relations are all related to more managerial/social aspects of digital.

In order to validate the structures derived based on the qualitative content-, and network analysis, the keywords of the 42 publications are used for science mapping. As this paper focuses on the constitutive characteristics of digital, the cluster digital is extracted from the overall network (compare Fig. 3) and nodes and edges are adapted as described in the methodology section of this paper. Zooming into the details of the digital cluster, the science mapping reveals the items of the digital cluster, showing a strong focus on IT-related topics and technical aspects of digital. Therefore, when considering features of digital, the focus was so far strongly influenced by technical and IT-related topics, while managerial and organizational aspects have been neglected.

Certain concepts or technologies like cyber-physical systems or Industry 4.0 are often mentioned when talking about digital. However, the nodes are not based on the same level of understanding, as one constitutes a technology and one is a strategy. Therefore, it is necessary to delimit digital from other concepts and to put it into perspective. In right graph of Fig. 3, the thickness and color of the edges represents the directed frequency of the connections, therefore the thicker an edge, the more frequent the connection, while the

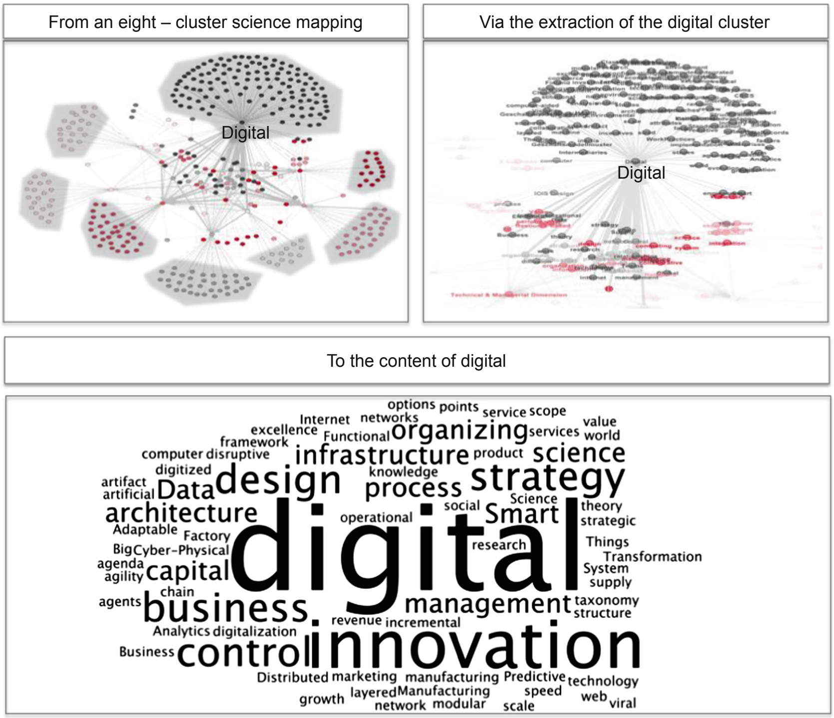


Fig. 3. From science mapping "Digital" to tag cloud content.

positioning of the different nodes represents the betweenness centrality, thus indicating how closely related the topics are to digital in contrast to other concepts mentioned. Thus, even though there might be nodes more closely related to the node digital, the ones highlighted in the graph are the ones most frequently named. The science mapping results in twelve different features that are strongly related to digital, namely computer-aided, 4.0, organizing, cyber-physical, analysis, architecture, data, big data, industry, adoption, review and exchange (compare Fig. 3, right side).

The categories derived from the qualitative and network analysis serve as a basis for defining the dimensions of digital, whereas the keywords from the science mapping constitute indicators for digital.

3 Results: Characteristics of the Digital Construct

The result of the descriptive and qualitative reflection of digital is an analysis scheme consisting of derived dimensions of digital. Based on the current state of research, it becomes apparent that a phase model is appropriate (Samulat 2017; Schönbohm and Egle 2017; Tilson et al. 2010; Yoo 2010; Yoo et al. 2010a, b). Four different phases are used. In order to make the phases clearly distinguishable, the phases will not only be numbered,

Table 2. Constituent characteristics of digital.

Characteristics / Phases			Digitization	Media convergence	Digitalization	Digital transformation
Definition			Conversion of analog signals to binary code	Fusion of various data formats	Synergetic use of technological innovations (result phase 1 and 2) on the broad base of society	Creation of a generative ecosystem based on technical and social components
Indicator			Technology	Decreasing cost for transaction- and reproduction	Services and transparency	Content
Relevant constituent characteristics	**Structuring dimensions**	Data generation	Real data		Real and simulation data	
		Enabler	Technology	Technology	Technology and analytics	Technology, analytics and service orientation
	Strategic dimensions	Control	Centralized		Decentralized	
		Expectation	Conversion of analog signals		Expand existing business models and develop new ideas	Digital Darwinism
		Technology needs to be …	Exploitation		Exploration	Exploitation
	Objective and performance dimensions	Supply chain focus	Internal	Product-centered	Internal, Dyadic	Chain, Network
		Performance measurement	Continuous, subjective	Continuous or to a certain point, subjective	n/A, objective	n/A, objective
	Amendment dimensions	Solution	Technological		Socio-**technical**	**Socio**-technical
		Amendment by	New technologies		New technologies, new corporate structure	Business model, customer centric
		Type of amendment	Process optimization	Product innovation	Process innovation, product optimization (business model optimization)	Business model innovation, process optimization, product innovation

but named according to their main focus, resulting in phase 1 – digitization, phase 2 – media convergence, phase 3 – digitalization and phase 4 – digital transformation.

Digitization describes the conversion of analogue signals into binary code as well as the electronic storage of coded data. **Media convergence** describes the fusion of various data formats using one platform or device. Phases 1 and 2 are characterized by technical changes. **Digitalization** describes the synergetic use of technical innovations resulting from the phases of digitization and media convergence, on the broad base of society, thus becoming a disruptive competitive advantage. Digitalization therefore relies on technical changes but is characterized by organizational changes. Beyond that, **digital transformation** describes the creation of a generative ecosystem of technical and social components, is a phenomenon mostly described in practice. Digital transformation thus primarily refers to a change of the business model, which has the changes of phases 1–3 as prerequisite.

The term **digital** thus encompasses all of the four phases described. Each phase can be described with a set of characteristics, foremost including a definition for each phase, followed by indicators that can be applied for identifying the different phases from an operational perspective at one glance. The dimensions of cluster 1, as identified in Fig. 2, have been broken down to the different characteristics of digital, thus operationalizing the nodes of the network. The relevant constituent characteristics can further be divided in four dimensions, with each having some sub-dimensions. The structuring dimension contains the data source (Stewart et al. 2004) and the enabling factors for the respective phase (Krickel 2015), thus indicating how the digital aspect is created. The strategic dimension focuses on control mechanisms, thus defining how control is exercised (Edwards et al. 2007), expectation management, thus describing what parties involved seek to achieve with their initiatives (Krickel 2015) and the operationalization of the technological components, which indicates what needs to be done with a technology (Faucet et al. 2011; Liu et al. 2016; Nagy 2006).

The objective and performance dimension analyses the breadth and direction of supply chain integration, as well as it defines if performance is measured continuously or to a certain point in time (Klein 2007; Gunasekaran and Kobu 2007; Chin et al. 2004; Kaplan 1990). The amendment dimension indicates which kind of changes are applied, i.e. technological, technical or social (Edwards et al. 2007; Yoo et al. 2010a, b), it analyses which kind of change leads to a change of the status quo (Kreutzer et al. 2017) and it reviews the nature of the amendment in terms of optimization and innovation (Weinreich 2016). The key aspects related to digital, as depicted in Fig. 3 mostly constitute examples of technologies or applications, thus they can be assigned as items for the different constituent characteristics. In a brief summary, the results of the literature analysis indicate, that even though digital seems to be quite versatile at first glance, in-depth analysis allows the derivation of a clear structure (compare Table 2, see annex 1 for category building).

4 Observations, Implications, and Conclusion

The first observation of this paper was that so far, a common understanding for digital has not been established. Even in purchasing and supply management, digital has not yet been treated explicitly, being described as underlying principle so far. By offering constitutive characteristics of digital in form of a maturity model, this paper tries on the one hand to close the conceptual dissent, on the other hand it offers specific and operationalized features. These might help to clarify theoretical discussion and can be used by practitioners to distinguish not only their current digital position, but also to disentangle the versatility by creating a common understanding.

The second observation was that concepts, applications and technologies are mixed up frequently. However, technology application and new socio-economic business making should clearly be distinguished from one another. A digital maturity model is proposed (Fig. 4) in order to provide a structure for allocating different phenomenon in the future. The emergence of information- and communication technologies (ICT) describes phase 1 – digitization (Use of digital technologies). The integration of these technologies is reflected in phase 2 – media convergence (Müller 2017). For the sake of completeness, Industry 4.0, as it has mentioned in connection with digital throughout this paper, will be delimited from digital as well: Industry 4.0 describes a future scenario of an interlinked and digitized production, meaning Internet-based automation and digitization of production processes are realized. Industry 4.0 thus describes a result of phases 1 and 2 in a production context (Huber 2013). Further stages relate to social use of digital, so a failure to social change in the long run leads to failure in developing a digital identity. Therefore, digitalization in today's understanding coins a socio-technical change, with technical implications that are being targeted already, and social implications that have mostly been neglected or underestimated so far.

The concluding remarks refer to the limitations: this work identified and derived a maturity model with underlying constituent features, but more conceptual and empirical work on how to operationalize and measure the dimensions and effects of digital is needed. As the research published on digital is growing exponentially, this research does only offer a snapshot of current relevant characteristics. Research on digital is not

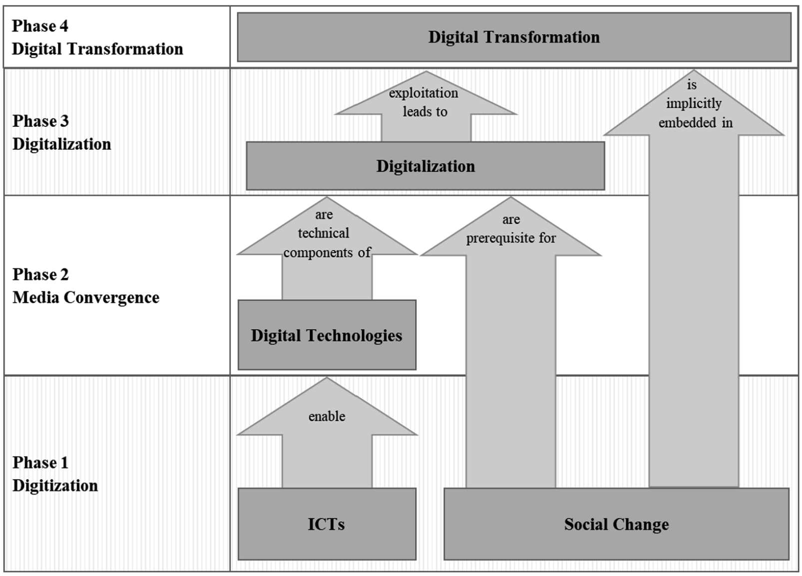
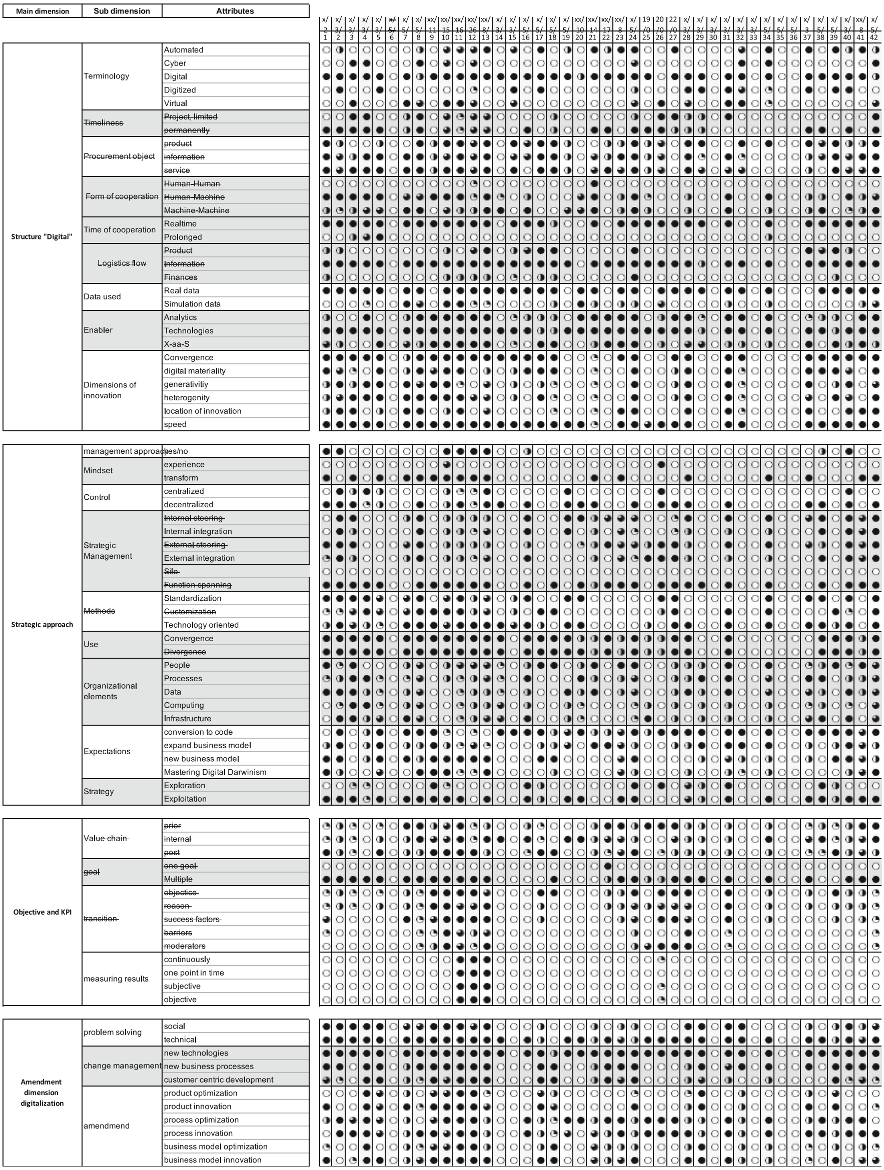


Fig. 4. Interrelation of digital components

only a technical one, functional perspectives such as purchasing and supply management have the chance and responsibility to contribute to how we understand digitalization/digital transformation in the future.

Annex 1: Category Building Based on Existing Journal Publications



Analysis of 42 papers based on abductive categories, categories which are crossed out have been identified as not relevant. Harvey balls indicate content of papers in qualitative 25% steps. Each colored quarter symbolizes more details on dimension in paper; a completely filled Harvey ball indicates topic is included explicitly in paper.

References

Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: Third International AAAI Conference on Weblogs and Social Media, pp. 361–362 (2009)

Bennett, S.: Control and the digital computer: the early years. In: 15th Triennial World Congress, Barcelona, Spain. IFAC (2002)

BMWi. Industrie 4.0 und Digitale Wirtschaft, Berlin (2015)

Borgatti, S.P.: Centrality and network flow. Soc. Netw. **27**, 55–71 (2005)

Bracht, U., Masurat, T.: The digital factory between vision and reality. Comput. Ind. **56**, 325–333 (2005)

Briner, R., Denyer, D.: Systematic Review and Evidence Synthesis as a Practice and Scholarship Tool. Oxford University Press, Oxford (2012)

Bullock, R.J., Batten, D.: It's just a phase we're going through: a review and synthesis of OD phase analysis. Group Organ. Stud. **10**, 383–412 (1985)

Capgemini Consulting. The Digital Advantage: How digital leaders outperform their peers in every industry [WWW Document]. MIT Sloan Management Review (2018). https://www.capgemini.com/wpcontent/uploads/2017/07/The_Digital_Advantage__How_Digital_Leaders_Outperform_their_Peers_in_Every_Industry.pdf. Accessed 9 July 2018

Chin, K.S., Rao, T., Leung, J.P.F., Tang, X.: A study on supply chain management practices The Hong Kong manufacturing perspective. Int. J. Phys. Distrib. Logist. Manage. **34**, 505–524 (2004)

Corbin, J., Strauss, A.L.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications, Newbury Park (2007)

Denyer, D., Tranfield, D.: Producing a systematic review. In: Buchanan, D., Bryman, A. (eds.) The Sage Handbook of Organizational Research Methods, pp. 671–689. SAGE Publications Ltd., London (2009)

Easterby-Smith, M., Thorpe, R., Jackson, P.: Management and Business Research, 5th edn. SAGE Publications Ltd., Los Angeles (2015)

Edwards, P., Jackson, S., Bowker, G., Knobel, C.: Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures, Michigan (2007)

Fisher, G., Aguinis, H.: Using theory elaboration to make theoretical advancements. Organ. Res. Methods **20**(3), 438–464 (2017)

Forrester, J.W.: Digital Computers as Information-Processing Systems (No. R-166-1), Digital Computer Laboratory, Massachusetts (1949)

Gartner. Digitalization - Gartner IT Glossary [WWW Document] (2018). https://www.gartner.com/it-glossary/digitalization/. Accessed 6 Jan 2018

Google. Digital [WWW Document] (2018). https://www.google.com/. Accessed 11 Apr 2018

Hassani, S.N.: Locating digital divides at home, work, and everywhere else (2006)

Henfridsson, O., Yoo, Y., Svahn, F.: Path creation in digital innovation: a multi-layered dialectics perspective. In: Association for Information Systems, Working Papers on Information Systems, pp. 9–20 (2009)

Huber, A.S.: Das Ziel Digital Enterprise: die professionelle digitale Abbildung von Produktentwicklung und Produktion. In: Sendler, U. (ed.) Industrie 4.0: Beherrschung Der Industriellen Komplexität Mit SysLM, pp. 111–125. Springer, Heidelberg (2013)

Kache, F., Seuring, S.: Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. Int. J. Oper. Prod. Manage. **37**(1), 10–36 (2017)

Knight, L., Tate, W., Matopoulos, A., Meehan, J., Salmi, A.: Breaking the mold: research process innovations in purchasing and supply management. J. Purchasing Supply Manage. **22**, 239–243 (2016)

KPMG. Advantage Digital [WWW Document] (2018). https://home.kpmg.com/uk/en/home/insights/2017/03/advantage-digital.html. Accessed 9 Jul 2018

Klein, R.: Customization and real time information access in integrated eBusiness supply chain relationships. J. Oper. Manage. **25**, 1366–1381 (2007)

Kreutzer, R.T., Neugebauer, T., Pattloch, A.: Digital Business Leadership Digitale Transformation - Geschäftsmodell-Innovation - Agile Organisation - Change Management. Springer, Berlin (2017)

Kreutzer, R.T., Land, K.-H.: Digital Darwinism Branding and Business Models in Jeopardy. Springer, Heidelberg (2015)

Krickel, F.: Zukunftsorientierte Unternehmenssteuerung in der Energiewirtschaft, Zukunftsorientierte Unternehmenssteuerung in der Energiewirtschaft. Springer, Wiesbaden (2015)

Liu, Z., Prajogo, D., Oke, A.: Supply chain technologies: linking adoption, utilization, and performance. J. Supply Chain Manage. **52**, 22–41 (2016)

Mayring, P.: Qualitative inhaltsanalyse. In: Flick, U., von Kardoff, E., Keupp, H., von Rosenstiel, L., Wolff, S. (eds.) Handbuch Qualitative Forschung: Grundlagen, Konzepte, Methoden Und Anwendungen, pp. 209–213. Beltz - Psychologie Verl. Union, München (1991)

Mayring, P.: Qualitative Inhaltsanalyse. Handbuch Qualitative Forschung in der Psychologie **1**, 601–613 (2010)

Miles, M.B., Huberman, A.M.: Qualitative Data Analysis: An Expanded Sourcebook, 2nd edn. Sage Publications, Thousand Oaks (1994)

Müller, H.-E.: Digitale Technologien [WWW Document]. Gabler Wirtschaftslexikon (2017). https://wirtschaftslexikon.gabler.de/definition/digitale-technologien-54127. Accessed 25 Sept 2018

Newman, M.E.J.: Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Phys. Rev. E **64**, 1–7 (2001)

Okamoto, K., Chen, W., Li, X.-Y.: Ranking of closeness centrality for large-scale social networks. In: Preparata, F., Wu, X., Yin, J. (eds.) Proceedings of the 2nd Annual International Workshop on Frontiers in Algorithmics, pp. 186–195. Springer, Berlin (2008)

Pagani, M., Pardo, C.: The impact of digital technology on relationships in a business network. Ind. Mark. Manage. **67**, 185–192 (2017)

Patnayakuni, N.: Firm performance impacts of digitally enabled supply chain integration capabilities. MIS Q. **30**, 225–246 (2006)

Pescher, J.: Change Management: Taxonomie und Erfolgsauswirkungen. Gabler Verlag, Wiesbaden (2009)

Poole, M.S.: Central issues in the study of change and innovation. In: Poole, M.S., Van de Ven, A.H. (eds.) Handbook of Organizational Change and Innovation, pp. 3–31. Oxford University Press, Oxford (2004)

Pratt, M.G., Rockmann, K.W., Kaufmann, J.B.: Constructing professional identity: the role of work and identity learning cycles in the customization of identity among medical residents. Acad. Manage. J. **49**, 235–262 (2006)

Pratt, M.K., Rouse, M.: What is ICT [WWW Document]. WhatIs (2005). https://searchcio.techtarget.com/definition/ICT-information-and-communications-technology-or-technologies. Accessed 25 Sept 2018

Samulat, P.: Die Digitalisierung der Welt: Wie das Industrielle Internet der Dinge aus Produkten Services macht. Springer Fachmedien, Wiesbaden (2017)

Schentler, P., Klein, A.: Einkaufscontrolling: Instrumente und Kennzahlen für einen höheren Wertbeitrag des Einkaufs. Haufe Lexware, Freiburg im Breisgau (2016)

Spiggle, S.: Analysis and interpretation of qualitative data in consumer research. J. Consum. Res. **21**, 491–503 (1994)

Srai, J.S., Lorentz, H.: Developing design principles for the digitalisation of purchasing and supply management. J. Purchasing Supply Manage. **25**(1), 78–98 (2019)

Stewart, D., Ewing, M., Mather, D.: e-Audience estimation: modelling the spread of viral advertising using branching theory. Institute for Operations Research and the Management Sciences Annual Meeting (2004)

Tilson, D.A., Lyytinen, K., Sorensen, C.: Desperately seeking the infrastructure in IS research: conceptualization of “digital convergence” as co-evolution of social and technical infrastructures. In: 43rd Hawaii International Conference on System Sciences, pp. 1–10. IEEE (2010)

Tilson, D.A., Lyytinen, K., Sørensen, C.: Research commentary—digital infrastructures: the missing IS research agenda. Inf. Syst. Res. **21**, 748–759 (2010)

Tilson, D.A.: The interrelationships between technical standards and industry structures: actor-network based case studies of the mobile wireless and television industries in the US and the UK. Weatherhead School of Management, Case Western Reserve University (2008)

Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence-informed management knowledge by means of systematic review* introduction: the need for an evidence-informed approach. Br. J. Manag. **14**, 207–222 (2003)

Vasudevan, R.K., Ziatdinov, M., Chen, C., Kalinin, S.V.: Analysis of citation networks as a new tool for scientific research. MRS Bull. **41**, 1009–1016 (2016)

Verband der Hochschullehrer für Betriebswirtschaft e.V.: Gesamtliste, 2018. VHB Journalqualität. http://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/gesamtliste/. Accessed 25 June 25 2018

Weinreich, U.: Lean Digitization. Springer, Heidelberg (2016)

Yoo, Y.: Computing in everyday life: a call for research on experiential computing. MIS Q. **34**, 213–231 (2010)

Yoo, Y., Boland, R., Lyytinen, K., Majchrzak, A.: Organizing for innovation in the digitized world. Organ. Sci. **23**, 1398–1408 (2012)

Yoo, Y., Henfridsson, O., Lyytinen, K.: Research commentary —the new organizing logic of digital innovation: an agenda for information systems research. Inf. Syst. Res. **21**, 724–735 (2010a)

Yoo, Y., Lyytinen, K., Boland, R., Berente, N.: The next wave of digital innovation: opportunities and challenges: a report on the research workshop “digital challenges in innovation research. SSRN Electron. J. 1–37 (2010b)

Zhang, L., Glänzel, W., Liang, L.: Tracing the role of individual journals in a cross-citation network based on different indicators. Scientometrics **81**, 821–838 (2009)

Zhang, L., Janssens, F., Liang, L., Glänzel, W.: Journal cross-citation analysis for validation and improvement of journal-based subject classification in bibliometric research. Scientometrics **82**, 687–706 (2010)

Zhu, K., Dong, S., Xu, S.X., Kraemer, K.L.: Innovation diffusion in global contexts: determinants of post-adoption digital transformation of European companies. Eur. J. Inf. Syst. **15**, 601–616 (2006)

Digital Forwarders
A Market Oriented Taxonomy

Ralf Elbert(✉) and Michael Gleser

Department of Law and Economics, Chair of Management and Logistics,
TU Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany
{elbert,gleser}@log.tu-darmstadt.de

Abstract. Digital forwarders are a rather new phenomenon in the logistics market. They claim benefits for shippers and freight carriers alike and threaten established players in the market. As a lot of digital forwarders are still situated in a start-up phase, the phenomenon has not yet arrived thoroughly in scientific discussions. Digital forwarders are not necessarily alike each other but try to cover various market segments. To enable a profound analysis of digital forwarders, this paper presents an empirically based taxonomy to serve as a starting point for both theoretical and practical discussions. The taxonomy was created through an iterative process by analyzing major digital forwarders in the market. The taxonomy consists of three major types within the digital forwarding market, being international full service, direct contract trucking and niche digital forwarder.

Keywords: Digitization · Freight forwarder · Taxonomy · Platform · Logistics · Disruptive innovation

1 Introduction

After the banking and retailing sector, digitization starts to hit the logistics sector as well (seen by growing interest in research by major logistics association [5]). Especially the availability of data in the areas of IoT and Industry 4.0 give rise for data driven business models [39]. New disruptive business models for the forwarding sector might come from start-up competitors as it has been the case for other industries like retail and entertainment [7]. The logistics sector has recently experienced huge inflow of risk capital to fund large numbers of start-ups [9,29,36]. With this increased inflow of risk capital chances for the establishment of innovative and disruptive business models like digital forwarders might take place. Reactions among established logistics players are twofold. Some see the development as a major threat to their businesses while others have a confident outlook due to their own investments in digitization [17,19,32].

A forwarders business is essentially securing the transportation of goods from a sender to a receiver either using own transportation capacity or making use of external freight carriers [31]. The business of a forwarder is highly information

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 19–31, 2019.
https://doi.org/10.1007/978-3-030-29821-0_2

based, be it for transport capacity acquisition or transport coordination [27]. This data oriented focus of a forwarder's business gives opportunity for digital platforms or algorithm based business models attacking established players in the market through disruptive innovation [16]. This disruption has taken place in the taxi business by Uber Inc. The business model of Uber may be applied to various other personalized services [34,35]. Uber Inc. has recently started to transfer their knowledge of platform based business models to the transportation sector through Uber Freight Inc. [38]. Just like Uber Inc. for the taxi business, Uber Freight Inc. uses an algorithm based platform approach to match transportation capacity demand and supply (both for the short and long term) and giving opportunity for truckers to directly contract with the platform, essentially allowing for a one truck company. Other competitors in the market for digital forwarding services are FreightHub [13], Saladoo [33], FlexPort [12] or Convoy [8].

Research in the area of electronic freight exchanges has long been established, regarding operations, structures and applications [3,6,20]. Electronic freight exchanges connect transport capacity demand and supply. The business model of digital forwarders seems similar to that of electronic freight exchanges as it connects carriers and shippers. Despite the similarity, research focus has not yet been directed towards digital forwarders, although the topic has found its way into professional news and grey literature [4,11,24]. The variety of digital logistics services offered in the forwarding sector makes it difficult to apply universal concepts.

2 Fundamentals and Development of Digital Forwarders

The business of a freight forwarder is to ensure the transportation of goods for shippers using either own transport capacity or making use of external freight carriers[1]. A multitude of different forwarder types have established themselves over time. Specialized forwarders for different modes (sea, air, road), different regional focus and different types of goods are just a few of the potential separations one can draw among them [30, p. 309].

Within this multitude of different freight forwarder types so called digital forwarders a beginning to emerge. A look into literature reveals, that there is little dedicated research geared towards the field of digital forwarders so far. The thematically closest work is a conference proceedings paper from 2018 analyzing the logistics start-up market in general [21]. Despite the lack of research, two definitions of digital forwarders have been identified. Stölzle et al. focus their definition of digital forwarders on automation potentials [18]. For them, digital forwarders are online platforms without own assets that calculate algorithm based routes and provide instantly quoted prices for requested lanes. All operational processes are hereby settled through the online platform. A similar definition is given by Dietrich and Fiege [10]. According to them, a digital forwarder offers

[1] The corresponding legal framework and definitions can be found in §453-466 HGB for the German market and Title 49 U.S.C. §13102 for the US market.

the same functional spectrum as a classical forwarder (e.g. transport organization, disposition and track and trace) and uses a digital platform to handle all needed processes and document exchanges through an online platform. Both definitions state that digital forwarders solely rely on external transporters and do not operate own transportation capacity. In contrast to legacy forwarders, digital forwarders are algorithm-driven and offer (mostly) a digital only interface to their service to allow for minimal staff intervention aside from trouble-shooting, while legacy forwarders still offer a multitude of different analogous contact options like phone or fax as well with the corresponding cost. As the market for digital forwarders is diverse and developing, a definite definition is yet difficult to give. The disruptive potential of digital forwarders will continue pressuring established players to focus investments at their digital competence as shown by a study of Oláh et al. shows [28]. This would essentially push established forwarders to adopt business practices of digital forwarders.

Digital forwarders have not yet been assessed from an empirical perspective, so we see clear need for such a contribution to the topic. As current definitions of digital forwarders diverge from market developments, we assume that there are multiple forms of digital forwarders, that can be separated by certain criteria. The research question we want to answer is therefore:

What Types of Digital Forwarders Can Be Derived from Empirical Analysis?

Stuart et al. [37] have developed a 5 step process model for empirical case studies which we will incorporate to answer the question at hand. After developing an explicit research question, the instrument for answering the question will be developed in the next section. Before applying the developed method the process for data gathering will be described. The final data analysis and development of a taxonomy round off the 5 stages.

3 Method Description for Empirical Taxonomy Creation

Classifications within the logistics sector have been conducted with various focus and width. Wen et al. classified air cargo forwarders from a service and competition perspective [41]. A more general approach classifying logistics service providers has been given by Lai et al. with an empirical perspective on the issue [23]. A classification with focus on information technology strategy for the Chinese market has been given by Lai, Zhao and Wang [22].

If there is little theoretical background for a certain field of research, building of such theory is a prerequisite for further practical research [15]. The building of a taxonomy based on empirical data can give such a theoretical foundation as shown by Gimpel et al. for the FinTech sector [14]. They based their taxonomy on the procedures described by Nickerson et al. and followed an iterative process in creating their taxonomy [25]. We have chosen to follow the approach conducted by Gimpel et al. as the FinTech sector has a digital first approach similar to the one followed by digital forwarders and therefore the transfer of concepts and procedures in creating a taxonomy for digital forwarders can be derived from the work of Gimpel et al.

The taxonomy creation process as used and described by Nickerson et al. follows an iterative approach [25, 26]. According to Nickerson et al. a taxonomy should be concise, sufficiently inclusive, comprehensive and extensible. In order to use the taxonomy process for the task at hand, it was adjusted to simplify the process at hand. Figure 1 gives an overview of the different steps of the followed taxonomy creation process.

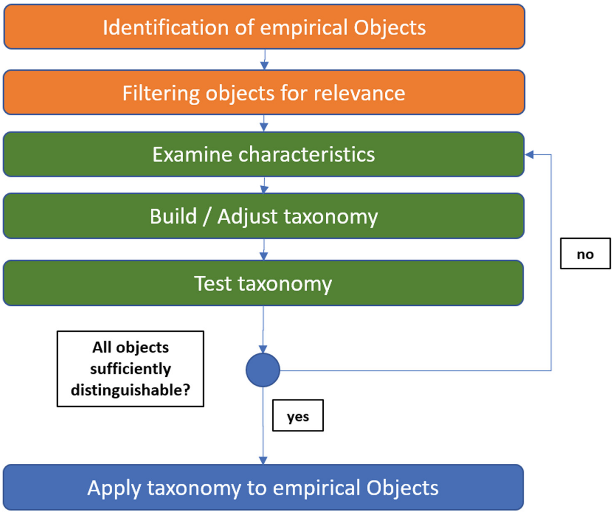


Fig. 1. Adjusted taxonomy creation process after Nickerson et al. (2009)

The adjusted taxonomy creation process involves three major phases as displayed in Fig. 1. As a first step, empirical objects need to be identified to be included in the process. As the objects identified might be irrelevant a pre-selection of matching objects needs to be conducted to increase homogeneity and relevance. After identifying relevant empirical objects to be included in the process these objects need to be analyzed for characteristics. These characteristics are the basis for a first taxonomy which is tested against the empirical objects at hand. As long as the taxonomy does not sufficiently separate the objects at hand the identification of relevant characteristics need to be further iterated. Once a stable taxonomy has been identified it can be used to group different empirical objects in it, giving room for further analysis.

4 Taxonomy Creation Process

4.1 Identification of Empirical Objects

The companies used in the analysis were identified using web searches at Google for the terms *digital forwarder*, *eForwarder* and the german term *digitaler Spediteur*.

Using a web search provides a trade-off between relevance and precision as analyzed by Weare and Lin [40]. As we aim for an exploratory view of recent developments we deemed relevance more important and a web search therefore suitable. Results were enhanced through companies from professional news. To determine which companies to include we oriented ourselves on the definitions by Stölzle [18] and Dietrich and Fiege [10]. As they limit their definitions to businesses with a similar service spectrum like classical forwarders, we deemed it justifiable to broaden the scope. First of all we included self-proclaimed digital forwarders that don't necessarily fit the definition like Uber Freight. Second we included FreightHub [13] which holds own transport capacity after the purchase of an Asian sea forwarder and would therefore not fit the definition fully. In total 21 companies that provide digital forwarding services were included in the analysis. The list of digital forwarders, their names and web addresses can be found in the appendix.

4.2 Iterative Identification of Characteristics

In order to identify the characteristics, we have chosen an iterative process. Using several iterations allows an analysis of the topic without the need for a prior absolute concept. As we aim for a qualitative analysis of the topic we deemed this sufficient. The basis for a first view on digital forwarders is derived from the approach of Bierwirth et al. who have used a similar view for analyzing electronic freight marketplaces [2]. The vertical perspective takes a look at the different relationships of digital forwarders towards their customers and suppliers. The horizontal perspective takes a look at the competitors of the digital forwarders, which in this case might be mainly legacy forwarders. To include an environmental and internal view, a third perspective was added taking a look on the internal business model as well as the technological environment. Figure 2 visualizes the initial perspective on the digital forwarder.

The basis for the analysis of the companies are publicly available information, especially the companies' websites. The self-proclaimed service spectrum and process organization is therefore the basis of the following analysis. Third party information were excluded to avoid bias. This process gives focus to an empirical view on the issue, leaving out considerations from literature, although additional triangulation through a literature analysis for the different identified characteristics could be a next step for result verification.

For each iteration one dimension of the initial view was taken as the basis for identification of characteristics. For each of these views the websites of the digital forwarders were analyzed for relevant and accessible information. Hereby a list was constructed to include all possible characteristics. For a thorough analysis of the empirical objects, only characteristics were considered for the final analysis that all websites of the involved digital forwarders provided. As an example, the profit model, although interesting for the analysis of digital forwarders, was only provided by a minority of the companies involved (if at all).

A main problem is when to stop the iteration process and deem it sufficient, as new characteristics might contribute to the taxonomy. As subjectivity can't be completely ruled out in this process step, we stopped when we found no further

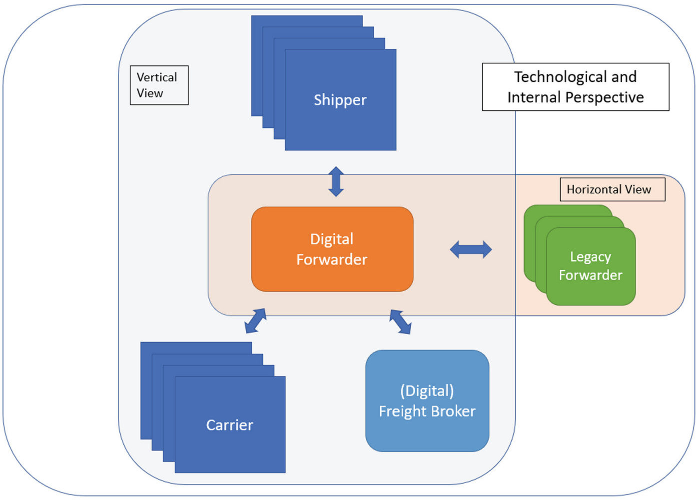


Fig. 2. Perspectives digital forwarder altered after Bierwirth et al. [2]

Table 1. Resulting dimensions and characteristics after iterative analysis

<table>
<tr><td colspan="5">Customer base</td></tr>
<tr><td colspan="3">Shippers</td><td colspan="2">Other forwarders</td></tr>
<tr><td colspan="5">Freight Capacity Supply</td></tr>
<tr><td colspan="2">Direct Carrier Registration</td><td colspan="2">Established Freight Network</td><td>Own Capacity</td></tr>
<tr><td colspan="5">Access Model for Customers</td></tr>
<tr><td colspan="3">Direct Web-Access</td><td colspan="2">Contract Logistics</td></tr>
<tr><td colspan="5">Contract Relationship</td></tr>
<tr><td colspan="3">Spot-based</td><td colspan="2">Recurring</td></tr>
<tr><td colspan="5">Quoting</td></tr>
<tr><td colspan="3">Instant-Quote</td><td colspan="2">Not explicitly stated</td></tr>
<tr><td colspan="5">Areal Focus</td></tr>
<tr><td colspan="3">International</td><td colspan="2">Regional</td></tr>
<tr><td colspan="5">Modal Focus</td></tr>
<tr><td>Sea</td><td>Air</td><td>Rail</td><td>Road</td><td>Storage</td></tr>
</table>

discriminating characteristics. In total 7 different characteristics were identified to distinguish the companies analyzed. The resulting list of characteristics is shown in Table 1.

4.2.1 Customer Base and Freight Capacity Supply

This dimension deals with the vertical perspective of the digital forwarder. Customers can either be shippers or other forwarders. While shippers are using digital forwarders to ship their freight, other forwarders can use platforms that offer the service to get rid of overcapacity or use the platform to obtain additional transportation demand. This model of cooperation among forwarders can lead to increased efficiency in the market at the cost of a co-opetition scenario (see Bengtson and Kock [1]).

When it comes to freight capacity supply the options were multitude. Some of the digital forwarders offered the possibility to directly register as a carrier on their website (e.g. No. 1, 10–12, 15–21) and start operating almost instantly. The most prominent companies representing this type of digital forwarder are Uber Freight and Convoy. Maintaining a classical freight network like established logistics service providers were followed by some of the digital forwarders as well (e.g. No. 2–10, 13, 14, 19). Another interesting characteristics by the digital forwarder Cargonexx was operations of own transportation capacity, contrasting the definition of Stölzle et al. [18].

4.2.2 Access Model, Contractual Focus and Quoting

Access to the digital forwarder can be twofold. Most of the forwarders analyzed offer a direct registration on their website and allow for use of their online platform (e.g. No. 3, 4, 7, 9, 10–15, 17–21). A lot of digital forwarders aim at acquiring contract logistics relationships which they fulfill using their digital platform (e.g. No. 2–10, 13, 16). Both models are closely related to the focus on the length of the service offered. Some of the forwarders explicitly aim at fulfilling spot-contracts (e.g. No. 3, 7, 9, 11–15, 17, 19–21).

A major characteristic that differentiates digital forwarders from each other and digital forwarders from legacy forwarders is their quoting system, essentially how to calculate prices for the service offered. Some digital forwarders claim to give instant quotes for requested lanes with instant booking (e.g. No. 1, 4, 6–8, 10, 13, 16, 17, 19–21). This is a major advance over legacy forwarders and highly dependent on sophisticated algorithms.

4.2.3 Areal and Modal Focus

A distinguishing characteristic of digital forwarders is their focus of service area. An international focus involves import and export processes. A local environment in this case means that it doesn't involve processes that cross legal borders which is the case in doing import or export.

On a modal focus, the four different modes that digital forwarders can serve are sea, air, rail and road. Additionally services in warehousing might be offered. The different modal offerings are a distinguishing characteristics, as some of the digital forwarders (e.g. No. 15–20 for trucking) only focus on one mode and try to be most efficient in it avoiding inter-modal transfers.

4.3 Suggested Taxonomy for Digital Forwarders

For identifying the actual types for the taxonomy the different dimensions were analyzed for their occurrence in combinations with other dimensions, which led to the identification of two main types. Direct contract trucking digital forwarder have a modal focus on trucking with direct carrier registration over a web interface and instant quoting. The full service international digital forwarder offers various transport modes with transportation capacity acquisition similar to that of classical forwarder (e.g. through an established network or the use of electronic freight brokers) and a focus on contract logistics and long term relationships. The remaining companies offered services mostly for a niche over a digital interface. Although their characteristics are heterogeneous they were grouped as niche digital forwarders as this best describes their market focus. Due to the size of the sample of 21 forwarders, further statistical analysis was deemed not reasonable.

4.3.1 Full Service International Digital Forwarder

The taxonomy type of full service international digital forwarders as shown in Fig. 3 offers a variety of modal services for international connections. Their focus lies on acquiring large portions of the forwarding market similar to legacy international logistics service providers.

Among the 21 companies identified in our analysis 9 were identified as international full service digital forwarders. The companies are BoxNBiz, Digital-Freight, FlexPort, FreightHub, Shippio, Shypple, Fleet, ZenCargo and the digital services of Kühne+Nagel KN FreightNet. The companies service offerings sound similar to each other, the differences should be assessed in a further study.

Customer Base				
Shippers	Other forwarders			
Freight Capacity Supply				
Direct Carrier Registration	Established Freight Network	Own Capacity		
Access Model for Customers				
Direct Web-Access	Contract Logistics			
Contract Relationship				
Spot-based	Recurring			
Quoting				
Instant-Quote	Not explicitly stated			
Areal Focus				
International	Regional			
Modal Focus				
Sea	Air	Rail	Road	Storage

Fig. 3. Full service international digital forwarder

4.3.2 Direct Contract Trucking Digital Forwarder

The taxonomy type of direct contract trucking digital forwarders as shown in Fig. 4 are such forwarders that focus solely on the highly flexible trucking business. Their business model is to offer instant-quoting to their customers while securing their transport capacity by directly contracting carriers through their online platform. Therefore their service is necessarily regional (when considering USA and EU).

Names within this taxonomy type are some of the companies that made the news claiming to be disruptive for the trucking sector. Among the 21 companies identified, 7 belong to this type. Their names are CargoMatic, CargoNexx, Convoy, Drive4Schenker, InstaFreight, MotherShip and Uber Freight.

Customer Base				
Shippers			Other forwarders	
Freight Capacity Supply				
Direct Carrier Registration		Established Freight Network		Own Capacity
Access Model for Customers				
Direct Web-Access			Contract Logistics	
Contract Relationship				
Spot-based			Recurring	
Quoting				
Instant-Quote			Not explicitly stated	
Areal Focus				
International			Regional	
Modal Focus				
Sea	Air	Rail	Road	Storage

Fig. 4. Direct contract trucking digital forwarder

4.3.3 Niche Digital Forwarder

The taxonomy type of niche digital forwarders as shown in Fig. 5 focuses on a certain mode of transportation or offer a digital interface to book loads within certain modes. Modes usually involved are sea, air and road. The essential difference to the trucking digital forwarders is that these companies need to establish a relationship with sea and air carriers prior to the start of operations as they need to digitally reserve and provide capacity. Companies that offer additional digital services were grouped in here as well.

Companies identified within the modal niche type are 5 out of the 21 companies. Their names are FreightOs, Saladoo, Twill, iContainers and Nyshex.

Customer Base				
Shippers			Other forwarders	
Freight Capacity Supply				
Direct Carrier Registration		Established Freight Network		Own Capacity
Access Model for Customers				
Direct Web-Access			Contract Logistics	
Contract Relationship				
Spot-based			Recurring	
Quoting				
Instant-Quote			Not explicitly stated	
Areal Focus				
International			Regional	
Modal Focus				
Sea	Air	Rail	Road	Storage

Fig. 5. Modal niche digital forwarder

5 Conclusion, Limitations and Outlook

The paper at hand has given an introduction to the field of digital forwarders, an emerging type of forwarder with focus on a digital platform based business model. Digital forwarders claim to provide benefit for both shippers and carriers by smoothing operations with digital administration, expanding capacity availability and increasing profits on all ends of the supply chain. Due to the novelty of the phenomenon, thorough research has not yet been conducted. Therefore a first taxonomy has been developed by assessing the market of digital forwarders and extracting characteristics suitable for classification. While technical features of the platforms have been left out (as done so by Gimpel et al. [14]) the focus has been set on the business aspects of these new competitors in the forwarding market. Next to the customer and capacity supply, contractual and business model issues have been covered as well as areal and modal focus of the platforms involved.

Through an iterative taxonomy creation process in total three different taxonomy types were identified to distinguish the sample of 21 companies. The identified types are international full service forwarders attacking the market of established international LSPs, direct contract trucking digital forwarders aiming towards increasing efficiency within the trucking sector and niche digital forwarders offering a multitude of different services for single aspects within the forwarding market.

While still in a start-up phase, some of the largest contestants in the digital forwarding sector aggressively aim to expand their market size and share which inevitably must come from established players in a settled market. The operational benefits of digital forwarders are clear to see, being lower operational costs due to a digital first approach, easy access and large capacity availability. The more these forwarders expand their modal capability into full service, the more traditional forwarders will come under pressure. With specialized digital forwarders catering specialized niches of the market we expect competition within the freight forwarding sector to increase.

There are several limitations of this study. First of all, only English and German service digital forwarders were included. Second of all, the selection of the forwarders was based on a web-search which might provide relevance at the cost of precision. And third of all, we only used the websites of the included digital forwarders as sources. The use of secondary sources, market share statistics or surveys of some form might be a welcome extension to precise results.

Further research should validate the identified types through qualitative approaches. Also the claims for service improvements at the digital forwarders should be empirically verified. For the direct contract trucking digital forwarders it would be interesting how these could potentially disrupt the transportation market.

The market of digital forwarding services is highly fascinating and dynamic. Research within this field needs to keep pace with the development of the start-ups themselves. This first market-based taxonomy aims to group existing digital forwarders into three types that are clearly distinguishable from each other.

Appendix

See Table 2.

Table 2. Considered digital forwarders

	Name	URL	Country	Type
1	BoxNBiz	https://www.boxnbiz.com/	India	Full Int. Service
2	Cargomatic	https://www.cargomatic.com/	USA	Direct Trucking
3	Cargonexx	https://www.cargonexx.com/	Germany	Direct Trucking
4	Convoy	https://convoy.com/	USA	Direct Trucking
5	DigitalFreight	https://www.digitalfreight.co.uk/	UK	Full Int. Service
6	Drive4Schenker	https://www.dbschenker.com/global/drive4schenker	Germany	Direct Trucking
7	Flexport	https://de.flexport.com/	Germany	Full Int. Service
8	FreightHub	https://freighthub.com/	Germany	Full Int. Service
9	Freightos	https://www.freightos.com/	Israel	Digital Niche
10	Instafreight	https://www.instafreight.de/	Germany	Direct Trucking
11	Mothership	https://www.mothership.com/	USA	Direct Trucking
12	Saloodo	https://www.saloodo.com/	Germany	Digital Niche
13	Shippio	https://www.shippio.io/	Japan	Full Int. Service
14	Shypple	https://shypple.com/	Netherlands	Full Int. Service
15	Fleet	https://www.tryfleet.com/	USA	Full Int. Service
16	Twill	https://www.twill.net/	Netherlands	Digital Niche
17	Uber Freight	https://www.uberfreight.com/	USA	Direct Trucking
18	Zencargo	https://www.zencargo.com/	UK	Full Int. Service
19	iContainers	https://www.icontainers.com/	Spain	Digital Niche
20	myKN	https://onlineservices.kuehne-nagel.com/	Switzerland	Full Int. Service
21	Nyshex	https://www.nyshex.com/en/	USA	Digital Niche

References

1. Bengtsson, M., Kock, S.: Coopetition in business networks-to cooperate and compete simultaneously. Industr. Market. Manag. **29**(5), 411–426 (2000). ISSN: 0019-8501
2. Bierwirth, C., Schneider, S., Kopfer, H.: Elektronische Transportmärkte Aufgaben, Entwicklungsstand und Gestaltungsoptionen. Wirtschaftsinformatik **44**(4), 335–344 (2002). ISSN: 1861-8936
3. Bohács, G., Frikker, I., Kovács, G.: Intermodal logistics processes supported by electronic freight and warehouse exchanges. Transp. Telecommun. **14**(3), 206–213 (2013)
4. Boston Consulting Group. The digital imperative in freight forwarding. https://www.bcg.com/de-de/publications/2018/digital-imperative-freight-orwarding.aspx. Accessed 29 May 2019

5. Bundesvereinigung Logistik (BVL): Positionspapier Digitalisierung. https://www.bcg.com/de-de/publications/2018/digital-imperative-freight-forwarding.aspx. Accessed 29 May 2019
6. Caplice, C.: Electronic markets for truckload transportation. Prod. Oper. Manage. **16**(4), 423–436 (2007)
7. Chesbrough, H.: Business model innovation: opportunities and barriers. Long Range Plan. **43**(2–3), 354–363 (2010)
8. Convoy. https://convoy.com/. Accessed 29 May 2019
9. Crunchbase - List of top Logistics Startups. https://www.crunchbase.com/hub/logistics-startups. Accessed 29 May 2019
10. Dietrich, A., Fiege, F.: Digitale Transformation des Speditionsgeschäfts umfasst mehr als Spedition 4.0. Wirtschaftsinformatik Manage. **9**(3), 36–45 (2017)
11. DVZ - Digitale Forwarder brauchen langen Atem. https://www.dvz.de/rubriken/logistik/detail/news/digital-forwarders-brauchen-langen-atem.html. Accessed 29 May 2019
12. FlexPort. https://de.flexport.com/. Accessed 29 May 2019
13. FreightHub. https://freighthub.com/de/. Accessed 29 May 2019
14. Gimpel, H., Rau, D., Röglinger, M.: Understanding in-Tech start-ups - a taxonomy of consumer-oriented service offerings. Electron. Markets **28**(3), 245–264 (2018). ISSN: 1422-8890
15. Gregor, S.: The nature of theory in information systems. MIS Q. **30**(3), 611–642 (2006). ISSN: 02767783
16. Hartmann, P.M., et al.: Capturing value from big data-a taxonomy of data-driven business models used by start-up firms. Int. J. Oper. Prod. Manag. **36**(10), 1382–1406 (2016)
17. Hofmann, E., Osterwalder, F.: Third-party logistics providers in the digital age: towards a new competitive arena? Logistics **1**(2) (2017). ISSN: 2305-6290
18. Kille, C., et al.: Digitalisierungswerkzeuge in der Logistik: Einsatzpotenziale, Reifegrad und Wertbeitrag, September 2018. ISBN: 9783736998797
19. Knowler, G.: Downplaying digital disruption: established forwarders confident they have what it takes to keep customers happy. J. Commer. **19**(9), 15–16 (2018). ISSN: 15423867
20. Kovács, G.: The structure, modules, services, and operational process of modern electronic freight and warehouse exchanges. Periodica Polytechnica Transp. Eng. **37**(1–2), 33–38 (2009)
21. Kurpjuweit, S., Wagner, S., Greil, T.: The rise of logistics startups and their impact on the logistics industry. In: 2018 25th EurOMA Conference (Budapest, Hungary), June 2018
22. Lai, F., Zhao, X., Wang, Q.: Taxonomy of information technology strategy and its impact on the performance of third-party logistics (3PL) in China. Int. J. Prod. Res. **45**(10), 2195–2218 (2007)
23. Lai, K.-h., Cheng, T.C.E., Yeung, A.C.L.: An empirical taxonomy for logistics service providers. Maritime Econ. Logist. **6**(3), 199–219 (2004). ISSN: 1479-294X
24. McKinsey and Company - Air-freight forwarders move forward into a digital future. https://www.mckinsey.com/industries/traveltransport-and-logistics/our-insights/air-freightforwarders-move-forward-into-a-digital-future. Accessed 29 May 2019
25. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. Eur. J. Inf. Syst. **22**(3), 336–359 (2013)

26. Nickerson, R., et al.: Taxonomy development in information systems: developing a taxonomy of mobile applications. In: HAL, Working Papers, January 2009
27. Nuzzolo, A., Coppola, P., Comi, A.: Freight transport modeling: review and future challenges. Int. J. Transp. Econ./Rivista internazionale di economia dei trasporti **40**(2), 151–181 (2013). ISSN: 03035247
28. Oláh, J., et al.: Information technology developments of logistics service providers in Hungary. Int. J. Logist. Res. Appl. **21**(3), 332–344 (2018)
29. Wyman, O.: Start-Ups rollen die Logistikbranche auf. https://www.oliverwyman.de/media-center/2017/Start-ups-rollen-Logistikbranche-auf.html. Accessed 29 May 2019
30. Pfohl, H.-C.: Logistiksysteme. Springer, Heidelberg (2018)
31. Pope, D.J., Thomchick, E.A.: U.S. Foreign freight forwarders and NVOCCs. Transp. J. **24**(3), 26–36 (1985). ISSN: 00411612, 2157328X
32. Produktion - Wie die Logistikbranche die Digitalisierung verschläft. https://www.produktion.de/trends-innovationen/wie-dielogistikbranchedie-digitalisierung-verschlaeft-103.html. Accessed 29 May 2019
33. Saladoo. https://www.saloodo.com/de/. Accessed 29 May 2019
34. Walker Smith, J.: The Uber-all economy of the future. Independent Rev. **20**(3), 383–390 (2016)
35. Sorescu, A.: Data-driven business model innovation. J. Prod. Innov. Manage. **34**(5), 691–696 (2017). ISSN: 07376782
36. Startup Region OWL - Alle 5 Tage wird ein Logistik Startup gegründet. https://www.startupregionowl.de/alle-5-tage-wirdein-logistik-startup-gegruendet/. Accessed 29 May 2019
37. Stuart, I., et al.: Effective case research in operations management: a process perspective. J. Oper. Manage. **20**(5), 419–433 (2002)
38. Uber Freight Inc. http://www.uberfreight.com. Accessed 29 May 2019
39. Waller, M.A., Fawcett, S.E.: Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. J. Bus. Logist. **34**(2), 77–84 (2013)
40. Weare, C., Lin, W.-Y.: Content analysis of the world wide web: opportunities and challenges. Soc. Sci. Comput. Rev. **18**(3), 272–292 (2000)
41. Wen, C.-H., Tsai, M.-C., Lin, C.-H.: Classification and competition analysis of air cargo logistics providers: the case of Taiwan's high-technology industry. J. Air Transp. Manage. **17**(2), 106–109 (2011). ISSN: 0969-6997

Human Role in Digital Logistics: Relevance of Intuition in Interacting with AI

Matthias Klumpp[1(✉)] and Caroline Ruiner[2]

[1] University of Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany
matthias.klumpp@uni-goettingen.de

[2] University of Trier, Universitätsring 15, 54296 Trier, Germany
ruiner@uni-trier.de

Abstract. Digital developments for logistics include many general and specific concepts as for example automation and Industry 4.0, Internet of Things (IoT), Physical Internet (PI) or Cyber-physical Systems (CPS). Overall, the human role in such settings will see profound changes – and many fears from workers are arising especially as there is no positive definition of new human work roles and expectations yet. We analyze the role of human intuition within an IoT and artificial intelligence application environment in logistics and supply chain processes and how it can be developed. Such a positive concept of increased efficiency by human-AI teams is an important cornerstone for digitalization as otherwise obstruction and fear may prevail with logistics and production workers.

Keywords: HCI · Intuition · Industry 4.0 · Physical Internet

1 Introduction

On first glance, human intuition and digitalization within a rational Internet of Things (IoT) environment in global value chains are antagonistic principles. On second thought, it can be argued that an interaction between the two approaches exists, with an important impact on decision quality and efficiency in logistics and Supply Chain Management (SCM). Such results are for example outlined by Turkulainen et al. [1] as they show and analyze the link of human motivation, work organization and leadership on the one hand and supply chain integration results and business implications on the other hand. Similarly, Alagaraja et al. [2] have exemplified the link between human resource and SCM advances for four US firms. This is again based on recent research advances e.g. by Carter et al. [3], conceptualizing intuition as a major human driving force in decision-making for logistics management. This new interest into soft factors in logistics and SCM is meeting a time when rapid changes as a result of hard technological factors within global value chains are virulent [see e.g. 4–6]. Autonomy concepts based on artificial intelligence (AI) application are the main driving forces of these developments [7–10]. A key question connected to these developments this concerns the future form and performance of human-computer interaction (HCI). In the past, working areas of robots and humans in value chains were separated and in the seldom cases of cooperation

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 32–44, 2019.
https://doi.org/10.1007/978-3-030-29821-0_3

e.g. in truck driving or CNC manufacturing, the roles were clearly defined: Human workers performed control and decision tasks, machines executed the mechanical tasks of transportation and production. That situation is changing as automation enters a new level of AI applications [11–14]. Robots and machines and more general both, assets (containers, transportation equipment) and products will be able to take advanced, self-reliant decisions without intervention, leaving the human workforce with only a superficial oversight role [15–17].

This is also motivating a new interest in the traditional “Asimov Laws” for robotics, dealing with safeguarding the interaction of humans on the one hand and robots and machines on the other hand. New momentum is put into this as increasingly all work areas also in logistics are featuring some sort of human-artificial interaction due to automation and AI implementation, regardless of blue- or white-collar job descriptions and tasks.

Consequently, the question of how humans arrive at decisions and which role intuition will play in cooperating with new technologies is of high importance for logistics [18–20]. So far, the link between intuition and digitalized work contexts has not been investigated in detail, although it can be expected that this will be of central relevance when it comes to HCI. To remedy this void it is the aim of this paper to analyze the role of human intuition within an Internet of Things (IoT) and AI application environment in logistics and supply chain processes and how it can be developed. For this, we refer to the concept of self-efficacy as this is a crucial prerequisite for the use of intuition and derive management implications.

The contribution of this paper is to establish a conceptual framework regarding the link from human intuition to IoT and AI application integration in logistics and SCM.

This paper is structured as follows: In the next section, the state of the art regarding the technical side (AI application concepts and implications, Physical Internet (PI) and IoT concepts in logistics) is reported. Following this, the human side is explored, hinting at important roles of trust, intuition and self-efficacy. After that, intuition as a new decision basis is explored further. Finally, cooperation options and implications regarding the specific role of intuition and trust in human-artificial interaction are outlined. The final section describes conclusions as well as follow-up research questions.

2 AI in SCM

An important group of professions in logistics and transportation are drivers and pilots for all transport modes on roads, rail, water and air transport networks. A number of measures is discussed and applied in order to improve efficiency and competitiveness in this work segment – with automation concepts in planning [21–24]. Other motivating factors to introduce automation and AI applications in transportation are safety and work conditions of drivers [25–28]. There is a significant potential of AI support in transportation, with e.g. truck cruise control systems as well as upcoming truck platooning concepts as lead examples for road transportation [29–31].

This is complemented by the vision of a completely automated supply chain, culminating in the PI. The PI was initially defined by Montreuil [32] as a logistics system where modular physical load packages are routed from source to destination through an automated network of hubs and spokes [see also 33 for further explanation]. Major elements of such a network do already exist for parcels, pallets, containers and swap bodies. Carriers for these types of loading units do optimize between alternative routes in their networks, e.g. through by-passing hubs. The PI is connected to IoT concepts as IoT describes communicating devices, including the option for local actions initiated by software agents. IoT technology is an important element of the PI, e.g. in cargo items independently determining alternative routes in case of congestion, or in signaling a potential quality loss in case of delays (e.g. in food transportation). The PI however is an alternative to a manually operated logistics network on a strategic decision level, with important consequences for all stakeholders involved. Such autonomous systems as the vision of the PI are enabled by latest technology developments in sensor applications as well as machine learning – combined allowing machine systems to accomplish complex tasks as for example autonomous driving. Such applications are determined to be "weak AI" applications [29, 34] as they are restricted to and focused on specific applications – in contrast to "strong AI" which would be able to solve a multitude and increasing range of tasks like humans do, then becoming "super-human". Machine learning is a cornerstone of AI applications as this concept allows a program to learn itself, leaving the former restrictions of coded actions behind: Whereas formerly every machine action had to be programmed in one way or another, software and therefore machines as e.g. robots are now able to learn from humans. For example the robotics innovative corporations Boston Dynamics (www.bostondynamics.com) from the US as well as Magazino from Germany (www.magazino.eu) have presented machines that are able to independently fulfill order picking tasks in intralogistics – and learn from humans regarding specific movements and hall layouts. In the Magazino case this represents also not individual but "swarm intelligence" learning as robots communicate among each other and each training effort is instantly shared with all machines within a group. For transportation and logistics there are already a large number of theoretical concepts and applications e.g. in the fields of traffic flow prediction and management [35, 36], transportation [37], production logistics [38, 39] or security and safety [40]. However, questions regarding the implementation of such automated concepts in logistics are mainly addressing the collaboration with humans and their acceptance of such systems.

3 Human Role in Digital Logistics

Human interaction with AI applications [10, 41] can be characterized by three areas of human resistance. Once a resistance area is overcome by the human actors, usually acceptance settles in [42, 43, 44], see Fig. 1. Three hurdles are connected to three functional areas of AI applications and represent an increasing level of human resistance in parallel with an increasing level of perceived personal intrusion for the human factors:

(a) AI competences: AI applications are acquiring competences in specific fields, from playing chess or go to forecasting weather and market demand. As separate competences, these AI competences are new for humans to be accustomed to but not

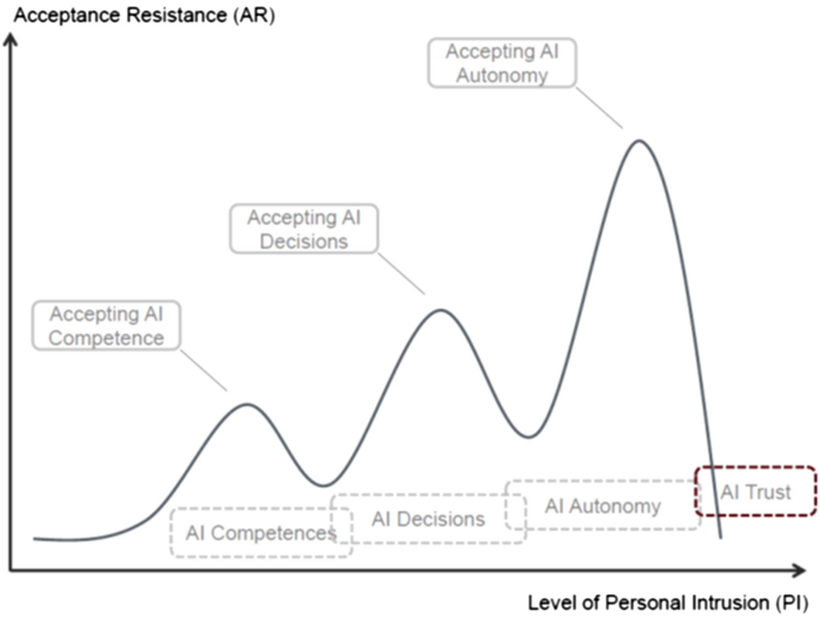


Fig. 1. Human acceptance model for artificial intelligence applications [44].

frightening and the resistance level towards them is low. For logistics and supply chain management, this includes for example the automated gearbox in trucks, automated routing and navigation systems as well as automated intralogistics applications like order retrieval and warehouse transportation systems. Such automated systems are characterized by the fact that any final decision is taken by a human.

(b) AI decisions: Increasingly, AI applications are providing management decision support, rising greater anxiety and resistance levels for humans. For example in an envisaged PI environment, where in AI applications calculate transport or stock ordering decisions [45, 46]. In such work and decision environments, humans are more anxious and possibly resistant because management decisions are addressed and human workers might be replaced or feel dictated and controlled. This sort of AI application is rising higher levels of rejection among humans, also requiring a longer period of adaption before acceptance can settle in as described by Weyer et al. [47].

(c) AI autonomy: Finally, AI applications are responsible for a bundle of decisions, constituting autonomous behavior as for example in largely automated manufacturing and shop floor environments or autonomous driving [38, 48]. In such cases, humans have to adopt a supervision role [49]. Such applications are in preparation for industrial production and logistics application areas, in production and warehousing [50, 51] or transportation. The highest level of resistance with humans emerges here, as they could feel excluded from day-to-day decisions and in many cases cannot really understand how AI applications arrive at specific decisions.

These resistance areas can be seen in connection with an increasing level of personal intrusion, possibly arriving at a new situation: The situation of trust with respect to AI applications, where humans are able to trustfully cooperate with AI applications [52]. This can be connected to the Turing test, where passing this test means that human actors are not able to distinguish between a human or artificial actor in blinded bilateral communication [53]. The stage of AI trust is a special form of passing the Turing test as it can be assumed that a human being may only be able to develop trust towards an AI application if an interaction-based evaluation can judge the collaboration partner to behave like a human being, i.e. trustworthy. This is a crucial and business-relevant form of trust between human beings and AI applications in logistics for a successful partnership. In addition, this is extending the traditional view of technology acceptance in the past [54], where application-specific trust and acknowledgement of human workers and customers was tested and analyzed. Long-term work situations like truck driving and machine handling with possible life-threatening situations are concerned with a required trust towards AI applications. Trust in turn can be seen as a major prerequisite for workers to develop intuition as it can only be trained and grown by trial and error experiences in practice, not in theory. Trust and intuition in this regard are interacting elements, e.g. intuition providing checks and balances on trust as well as trust supporting higher levels of intution or enacting intuitive decisions.

4 Intuition and Self-Efficacy

Recently, it has been argued that intuition might be able to complement rationality as an effective decision-making approach in organizations [55–57]. Intuition helps to cope with a wide range of critical decisions and is integral to successfully completing tasks that involve high complexity and short time horizons [58]. However, conceptualizations of intuition lack clarity so far. In general, intuition is differentiated in reliance on gut feelings (creative intuition) or in reliance on experiences (justified intuition) [59]. Adding to this discussion, Carter et al. [3] consider intuition as a major human driving force in decision-making for logistics management. On the basis of a qualitative content analysis of academic literature and interviews with SCM experts and quantitative testing with experts in supplier selection, they conceptualize intuition as a multidimensional construct consisting of the following dimensions: (a) Experience-based intuition implies that persons recognize parallels to past decisions in making the current decision and, thus, refer to knowledge that builds over time, (b) emotional processing means that affect or (positive and negative) feelings guide decisions and actions, and (c) automatic processing implies that persons quickly and almost instantly decide without awareness or knowledge of specific decision rules. Referring to their analyses, they argue that intuition is rather experience-based in situations with high time pressure, while emotional processing can be found in contexts with both, high time pressure and high information uncertainty. Against this background, obviously dual-processing theories (intuition vs. rationality) seem be too simplistic since different dimensions of intuition occur likewise.

Any approach to intuition has in common that intuitive judgments occur beneath the level of conscious awareness (i.e. they are tacit). The importance of intuition in working behavior has already been emphasized for top executives revealing that intuition was one of the skills used to guide their most important decisions [56, 60]. Khatri and Ng [61] surveyed senior managers of companies representing computer, banking, and utility industries in the United States and highlight that intuitive processes are in fact used in organizational decision making conceptualizing intuition as a form of expertise or distilled experience based on deep knowledge. In addition, Hogarth [62] argues that intuition is effective when a person is knowledgeable and experienced within a certain domain. The effective use of intuition has even been seen as critical in differentiating the more from less successful workers [58]. Especially for workers in logistics, intuition seems to be decisive in view of the requirement to cope with a wide range of critical decisions, highly complex tasks and short time horizons. This even increases through digitalization and tight networks.

In search for approaches to enhance intuition, Agor [60] reveals that physical and/or emotional tension and anxiety or fear are conceived as factors that impede the use of intuition, and in turn, positive feelings such as excitement lead to an enhanced sense of confidence in own judgments. Burke, Miller [63] argue that mentors, role models, and supervisors as well as working with diverse groups of people and learning about their decision making styles help to develop intuition. In this sense, intuition is based in implicit learning [64] and automaticity [65].

The use of intuition is more likely when people show high self-efficacy. Especially the Social Cognitive Theory [SCT; 66] accounts for the development and maintenance of self-efficacy across a broad range of knowledge, values, and associated behavior patterns [67]. An efficacy expectation is the conviction that a person can successfully execute the behavior required to produce certain outcomes. In this sense, self-efficacy means that coping behavior will be initiated, how much effort will be expended, and how long it will be sustained in the face of obstacles and aversive experiences [68]. Thus, self-efficacy regulates behavior, effort, and persistence over extended periods of time and is concerned with individual beliefs about ones capability to cope with certain and new situations, to use ones abilities to solve problems and tasks, e.g. whether they feel competent in using and learning how to use new technologies. In this respect, it can be expected that individuals with high self-efficacy are more proactive and confident in their decision-making, dealing with digital devices and show greater intuition.

In work contexts, especially following leverages are identified to support the worker's self-efficacy: (a) Performance accomplishments are based on personal mastery experiences in a sense that repeated success enhances self-efficacy, (b) vicarious experience means that seeing others performing activities and its consequences generate expectations that persons will improve and intensify their efforts as well, (c) verbal persuasion is quite common to influence human behavior, i.e. people are led through suggestions, and, finally, (d) emotional arousal affects perceived self-efficacy in coping with certain situations, i.e. removing dysfunctional fears.

So far, the link between intuition and self-efficacy has not been discussed yet although these two concepts show striking commonalities and provide approaches to further develop human intuition. As outlined, AI is developing fast within the supply chain and logistics management domain and there is a considerable body of literature concerning intuition and self-efficacy. However, both subjects are brought forward mainly independent of each other. This constitutes a major research gap, as obviously HCI concepts and existing business experience are testimony to the fact that the human factor is largely influencing technology implementation and AI efficiency, especially in logistics. In a prognosis perspective, research concepts combining these perspectives are needed urgently as mid- to long-term lead times can be expected until results are obtained and transferred into applicable concepts for increasing AI application effectiveness in logistics. Moreover, the role of intuition has already been emphasized for successful workers in other fields guiding important decisions in working life. In SCM, previous analyses concentrated on experts. Blue-collar workers which account for the largest proportion of employees in logistics have not been moved into the center yet. Finally, the concept of self-efficacy has been neglected so far, although it provides concrete approaches to further develop the use of intuition as basis for effective decisions in digitalized work contexts in logistics.

5 Relevance of Intuition in AI Application Settings

Human intuition plays an important role in cooperating with new technologies since workers mainly perform control and decision tasks supported by AI referring to improved analytics of increased data volume. Since human workers are in most cases not able to comprehend the data basis and algorithms that guided the AI suggestions, intuition is needed. This is because workers have to approve or disapprove of AI decision suggestions, without deeper understanding. Therefore, experience, “gut feeling” and other forms of intuition are important to arrive at efficient decisions. Because rejecting AI decision suggestions too often is not efficient – but missing the ones that would have to be rejected can be even more costly. This setting requires trust of workers on the decisions proposed, therefore intuition and trust are depending on each other. Intuition may help to doubt trust and re-examine unusual situations. On the other hand, trust also supports intuition in the form of trust in the own competences and experiences. Human workers shall not act as “blind” slaves of AI and machines, therefore trust and self-belief are crucial. Several aspects can be discussed regarding the relevance of intuition in collaborative HCI settings of logistics processes. Details are depicted in Fig. 2.

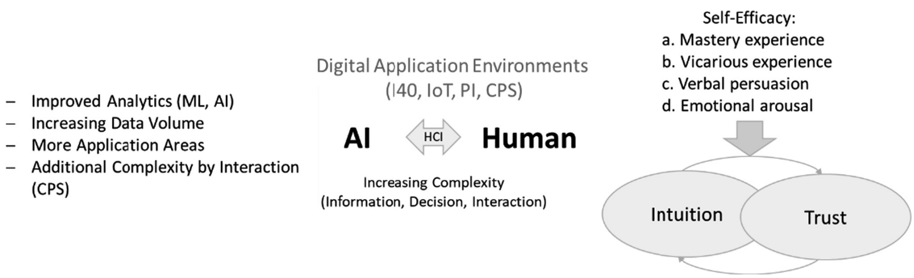


Fig. 2. AI and intuition interaction for human digital work in logistics.

We argued that self-efficacy is a crucial prerequisite for the application of intuition in digitalized work contexts in logistics processes. Thus, with increasing self-efficacy workers are more likely to use intuition. This might have at least two reasons: (1) an increased level of self-efficacy leads to more intuitive options for logistics decisions because of the underlying mastery and vicarious experiences as well as the consideration of peers' suggestions (verbal persuasion) and relaxed attitude (emotional arousal), and also (b) the ability to actually implement intuitive options for logistics decisions increases with a high level of self-efficacy. The worker's self-efficacy can be developed through competence development measures, as this would improve the use of intuition in complex decision settings. In this sense, competence management could enhance self-efficacy and, thus, the intuition of workers as critical prerequisite for successful collaboration of human workers and AI. This could have a high impact on the overall system performance in HCI logistics environments like e.g. in IoT or PI settings. Training and the resulting capabilities are important basic elements for human workers to deal with these digital work settings.

We propose that a beneficial human cooperation imperative as core design and evaluation element for AI systems. In this regard, we suggest to add a fourth law of beneficial human cooperation to the existing three Laws of Robotics. These laws were introduced by Isaac Asimov in his Runaround story of 1942: Whereas the three existing Asimov laws are directed only at the artificial entity like a robot, the fourth law should be directed at human-computer interaction. This is important as for example existing regulations like the directive 2006/42/EG as of May 17, 2006 for the European Union only address machines and artificial entities in order to promote safety of human users. Interaction aspects are usually not addressed in any safety or work regulations. The first law is stating that no robot is allowed to actively hurt a human being or to accept by inactivity that a human being is hurt. The second law according to Asimov is claiming that a robot always has to obey human orders – with the exception that this must not collide with the first law. The third law is proposing that a robot has to secure its own existence as long as this does not counteract laws number one and two [69, 70]. In addition to this, a new fourth law should address the question of HCI for example as follows: Robots are required to work together with humans if available and if not contradicting law number one and two. Actually, according to the priority sequence, this would become the new third law and the existing third robotics law would be placed fourth.

The named aspects are crucial elements of a comprehensive management concept for effective HCI. Such a concept of combined human-computer intuition-based management jointly applying intuition and rationality can be an important element for a competitive advantage regarding logistics and supply chain processes in a future IoT environment.

6 Discussion and Outlook

In outlining a conceptual framework regarding the influence of human intuition with regard to HCI in IoT and AI application environments, a link from human intuition to IoT and AI application integration in logistics and SCM is established. Moreover, we proposed implications as concept rules for enhancing intuition development in HCI within the logistics and supply chain domain. Possible implications and guidelines as competence management concept elements for enhancing work results in HCI are explained in detail. This connects to recent developments in AI and computer science itself, allowing AI applications to actively recognize trust by a human cooperation partner and react to that analysis result [71].

Regarding logistics workers, self-efficacy is expected to be decisive for the use of digital devices. Especially in these fields, we find new technologies and digitalized working contexts requiring the intuition of the workers, affecting their behavior, the cooperation with colleagues and interaction with supervisors. However, research is required to understand the role of human intuition within an IoT and AI application environment in logistics and supply chain processes. This is connecting to leadership and management theories such as transformational leadership giving workers the opportunity to make individual decisions in daily work processes, to further develop their competencies and to self-actualize [72]. Also the role of the worker's maturity would have to be reflected as addressed by other leadership theories such as Hersey and Blanchard [73].

To further investigate the role of human intuition in digitalized work settings in logistics and supply chain processes a qualitative study focusing on white- and blue-collar workers in logistics would help to explore the field. This is adding knowledge to the existing applications and explanatory models regarding human intuition – see for example [74–76]. Although, we focus on intuition and self-efficacy as the central determinants for the human-AI collaboration, further research should broaden the perspective and include other factors, e.g. professional competence, the design of digital devices and its perceived ease of use in everyday business as well. Not least, intuition and self-efficacy is connected to the element of leader support and other standard HR elements to be considered in future concept developments. Especially leader support and AI application acknowledgement towards workers within logistics processes are expected to be important inputs leading to possibly increased work results in logistics processes. This highlights the general focus of this paper that the human factor has to be analyzed and steered more closely and intensively especially in automatization and IoT context settings in logistics for successful results in the future.

Altogether, the future competitiveness and logistics performance will depend on the described factors regarding human intuition, motivation, and HCI. The challenge for logistics management will include the question of how to align and propagate talent and leadership management concepts with IoT and AI developments in the light of the important role of human intuition as highlighted in this paper.

References

1. Turkulainen, V., Roh, J., Whipple, J.M., Swink, M.: Managing internal supply chain integration: integration mechanisms and requirements. J. Bus. Logistics **38**(4), 290–309 (2017). https://doi.org/10.1111/jbl.12165

2. Alagaraja, M., Egan, T.M., Woodman, R.W.: Towards an understanding of HRD-lean strategy linkages: a multicase study of four U.S. based firms. Hum. Resour. Develop. Q. **28** (4), 553–592 (2017). https://doi.org/10.1002/hrdq.21292
3. Carter, C.R., Kaufmann, L., Wagner, C.M.: Reconceptualizing intuition in SCM. J. Bus. Logistics **38**(2), 80–95 (2017)
4. Hilger, T., Sahling, F., Tempelmeier, H.: Capacitated dynamic production and remanufacturing planning under demand and return uncertainty. OR Spectrum **38**, 849–876 (2016)
5. Bloemhof, J.M., van der Vorst, J.G., Bastl, M., Allaoui, H.: Sustainability assessment of food chain logistics. Int. J. Logistics Res. Appl. **18**(2), 101–117 (2015)
6. Fors, C., Kircher, K., Ahlström, C.: Interface design of eco-driving support systems–Truck drivers' preferences and behavioural compliance. Transp. Res. Part C **58**, 706–720 (2015)
7. Torabi, S.A., Hassini, E., Jeihoonian, M.: Fulfillment source allocation, inventory transshipment, and customer order transfer in e-tailing. Transp. Res. Part E: Logistics Transp. Rev. **79**, 128–144 (2015)
8. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015). https://doi.org/10.1038/nature14539
9. Gunsekaran, A., Ngai, E.W.T.: Expert systems and artificial intelligence in the 21st century logistics and SCM. Expert Syst. Appl. **41**(1), 1–4 (2014)
10. Lee, H.-D., Lee, B.-K., Kim, W.-S., Han, J.-S., Shin, K.-S., Han, C.-S.: Human-robot cooperation control based on a dynamic model of an upper limb exoskeleton for human power amplification. Mechatronics **24**(2), 168–176 (2014)
11. Knoll, D., Prüglmeier, M., Reinhart, G.: Predicting future inbound logistics processes using machine learning. Procedia CIRP **52**, 145–150 (2016)
12. Musa, A., Gunasekaran, A., Yusuf, Y., Abdelazim, A.: Embedded devices for supply chain applications: towards hardware integration of disparate technologies. Expert Syst. Appl. **41** (1), 137–155 (2014)
13. Zhang, S., Lee, C.K.M., Chan, H.K., Choy, K.L., Zhang, W.: Swarm intelligence applied in green logistics: a literature review. Eng. Appl. Artif. Intell. **37**, 154–169 (2014)
14. Wong, W.K., Leung, S.Y.S., Guo, Z.X., Zeng, X.H., Mok, P.Y.: Intelligent product cross-selling system with radio frequency identification technology for retailing. Int. J. Prod. Econ. **135**, 308–319 (2012)
15. Cantor, D.E.: Maximizing the potential of contemporary workplace monitoring: techno-cultural developments, transactive memory, and management planning. J. Bus. Logistics **37** (1), 18–25 (2016)
16. Crainic, T.G., Montreuil, B.: PI enabled hyperconnected city logistics. Transp. Res. Procedia **12**, 383–398 (2016)
17. Castelfranchi, C., Falcone, R.: Trust and control: a dialectic link. Appl. Artif. Intell. **14**(8), 799–823 (2000)
18. Gurkaynak, G., Yilmaz, I., Haksever, G.: Stifling artificial intelligence: human perils. Comput. Law Secur. Rev. **32**(5), 749–758 (2016)
19. Kim, T.G., Hornung, S., Rousseau, D.M.: Change-supportive employee behavior: antecedents and the moderating role of time. J. Manag. **37**(6), 1664–1693 (2011)
20. Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., Combs, B.: How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. Policy Sci. **9**(2), 127–152 (1978)
21. Gingerich, K., Maoh, H., Anderson, W.: Classifying the purpose of stopped truck events: an application of entropy to GPS data. Transp. Res. Part C **64**, 17–27 (2016)
22. Dondo, R., Cerdá, J.: The heterogeneous vehicle routing and truck scheduling problem in a multi-door cross-dock system. Comput. Chem. Eng. **76**, 42–62 (2015)

23. Phan, M.-H., Kim, K.H.: Negotiating truck arrival times among trucking companies and a container terminal. Transp. Res. Part E: Logistics Transp. Rev. **5**, 132–144 (2015)
24. Verma, M., Verter, V.: A lead-time based approach for planning rail-truck intermodal transportation of dangerous goods. Eur. J. Oper. Res. **202**, 696–706 (2010)
25. Chen, G.X., Sieber, W.K., Lincoln, J.E., Birdsey, J., Hitchcock, E.M., Nakata, A., Robinson, C.F., Collins, J.W., Sweeney, M.H.: NIOSH national survey of long-haul truck drivers: injury and safety. Accid. Anal. Prev. **85**(1), 66–72 (2015)
26. Pahukula, J., Hernandez, S., Unnikrishnan, A.: A time of day analysis of crashes involving large trucks in urban areas. Accid. Anal. Prev. **75**, 155–163 (2015)
27. Pattinson, W., Thompson, R.G.: Trucks and bikes: sharing the roads. Procedia-Social Behav. Sci. **125**, 251–261 (2014)
28. Khorashadi, A., Niemeier, D., Shankar, V., Mannering, F.: Differences in rural and urban driver-injury severities in accidents involving large-trucks: an exploratory analysis. Accid. Anal. Prev. **37**(5), 910–921 (2005)
29. Hengstler, M., Enkel, E., Duelli, S.: Applied artificial intelligence and trust – The case of autonomous vehicles and medical assistance devices. Technol. Forecast. Soc. Chang. **105**, 105–120 (2016)
30. Bernhart, W., Winterhoff, M., Hoyes, C., Chivukuly, V., Garrelfs, J., Jung, S., Galander, S.: Autonomous Driving. Roland Berger, Munich (2014)
31. Jamson, A.H., Merat, N., Carsten, O.M., Lai, F.C.: Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transp. Res. Part C: Emerg. Technol. **30**, 116–125 (2013)
32. Montreuil, B.: Towards a PI: meeting the global logistics sustainability grand challenge. Logistics Res. **3**(2), 71–87 (2011)
33. Ballot, E., Montreuil, B., Meller, R.: The PI – The Network of Logistics Networks. Programme de recherche et d'innovation dans les transports terrestres (PREDIT). La Documentation française, Paris (2014)
34. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs (2010)
35. Omrani, H.: Predicting travel mode of individuals by machine learning. Transp. Res. Procedia **10**, 840–849 (2015)
36. Julio, N., Giesen, R., Lizana, P.: Real-time prediction of bus travel speeds using traffic shockwaves and machine learning algorithms. Res. Transp. Econ. **59**, 250–257 (2015)
37. Wojtusiak, J., Warden, T., Herzog, O.: Machine learning in agent-based stochastic simulation: inferential theory and evaluation in transportation logistics. Comput. Math Appl. **64**, 3658–3665 (2012)
38. Wang, X., Tang, L.: A machine-learning based memetic algorithm for the multi-objective permutation flowshop scheduling problem. Comput. Oper. Res. **79**, 60–77 (2017)
39. Schuhmacher, J., Hummel, V.: Decentralized control of logistic processes in cyber-physical production systems at the example of ESB Logistics Learning Factory. Procedia CIRP **4**, 19–24 (2016)
40. Marucci-Wellman, H.R., Corns, H.L., Lehto, M.R.: Classifying injury narratives of large administrative databases for surveillance – A practical approach combining machine learning ensembles and human review. Accid. Anal. Prev. **98**, 359–371 (2017)
41. Bendoly, E.: Fit, bias, and enacted sensemaking in data visualization: frameworks for continuous development in operations and SCM analytics. J. Bus. Logistics **37**(1), 6–17 (2016)
42. Martínez-Torres, M., Díaz-Fernández, M., Toral, S., Barrero, F.: The moderating role of prior experience in technological acceptance models for ubiquitous computing services in urban environments. Technol. Forecast. Soc. Chang. **91**, 146–160 (2015)

43. Rousseau, D.M., Sitkin, S.B., Burt, R.S., Camerer, C.: Not so different after all: a cross discipline view of trust. Acad. Manag. Rev. **23**(3), 393–404 (1998)
44. Klumpp, M.: Automation and artificial intelligence in business logistics systems: human reactions and collaboration requirements. Int. J. Logistics Res. Appl. **21**(3), 224–242 (2018). https://doi.org/10.1080/13675567.2017.1384451
45. Phan, S., Ballot, E., Huang, G.Q., Montreuil, B.: PI and interconnected logistics servicves: research and applications. Int. J. Prod. Res. **55**(9), 2603–2609 (2017)
46. Sternberg, H., Norrman, A.: The PI – review, analysis and future research agenda. Int. J. Phys. Distrib. Logistics Manag. **47**(8), 736–762 (2017)
47. Weyer, J., Fink, R.D., Adelt, F.: Human-machine cooperation in smart cars: an empirical investigation of the loss-of-control thesis. Saf. Sci. **72**, 199–208 (2015)
48. Zhong, R.Y., Xu, C., Chen, C., Huang, G.Q.: Big data analytics for PI-based intelligent manufacturing shop floors. Int. J. Prod. Res. **55**(9), 2610–2621 (2017)
49. Rauffet, P., Lassalle, J., Leroy, B., Coppin, G., Chauvina, C.: The TAPAS project: facilitating cooperation in hybrid combat air patrols including autonomous UCAVs. Procedia Manuf. **3**, 974–981 (2015)
50. De Santis, R., Montanari, R., Vignali, G., Bottani, E.: An adapted ant colony optimization algorithm for the minimization of the travel distance of pickers in manual warehouses. Eur. J. Oper. Res. **267**(1), 120–137 (2017)
51. Li, B.-H., Hou, B.-C., Yu, W.-T., Lu, X.-B., Yang, C.-W.: Applications of artificial intelligence in intelligent manufacturing: A review. Front. Inf. Technol. Electron. Eng. **18**(1), 86–96 (2017)
52. Lahno, B.: On the emotional character of trust. Ethical Theory Moral Pract. **4**(2), 171–189 (2001)
53. Turing, A., Braithwaite, R., Jefferson, G., Newman, M.: Can automatic calculation machines be said to think? In: Copeland, B.J. (ed.) The Essential Turing – Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence and Artificial Life, pp. 487–505. Oxford University Press, Oxford (1952). 2004
54. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag. Sci. **46**(2), 186–204 (2000)
55. Calabretta, G., Gemser, G., Wijnberg, N.M.: The interplay between intuition and rationality in strategic decision making: a paradox perspective. Organ. Stud. **38**(3–4), 365–401 (2017)
56. Salas, E., Rosen, M.A., DiazGranados, D.: Expertise-based intuition and decision making in organizations. J. Manag. **36**(4), 941–973 (2010)
57. Sinclair, M., Ashkanasy, N.M.: Intuition. Myth or a decision-making tool? Manag. Learn. **36** (3), 353–370 (2005)
58. Dane, E., Pratt, M.G.: Exploring intuition and its role in managerial decision making. Acad. Manag. Rev. **32**(1), 33–54 (2007)
59. Stanczyk, A., Foerstl, K., Busse, C., Blome, C.: Global sourcing decision-making processes: politics, intuition, and procedural rationality. J. Bus. Logistics Res. **36**(2), 160–181 (2015)
60. Agor, W.H.: The logic of intuition: how top executives make important decisions. Org. Dyn. **14**(1), 5–18 (1986)
61. Khatri, N., Ng, H.A.: The role of intuition in strategic decision making. Hum. Relat. **53**(1), 57–86 (2000)
62. Hogarth, R.M.: Educating Intuition. University of Chicago Press, Chicago (2001)
63. Burke, L.A., Miller, M.K.: Taking the mystery out of intuitive decision making. Acad. Manag. Exec. **13**(1), 91–99 (1999)
64. Lieberman, M.D.: Intuition: a social cognitive neuroscience approach. Psychol. Bull. **126**(1), 109–137 (2000)

65. Moors, A., De Houwer, J.: Automaticity: a theoretical and conceptual analysis. Psychol. Bull. **132**(2), 297–326 (2006)
66. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice Hall, Upper Saddle River (1986)
67. Bandura, A.: Self-efficacy: the exercise of control. W. H. Freeman & Co, New York (1997)
68. Bandura, A.: Self-efficacy: toward a unifying theory of behavioral change. Psychol. Rev. **84** (2), 191–215 (1977)
69. Trentesaux, D., Rault, R.: Designing ethical cyber-physical industrial systems. IFAC Papers **50**(1), 14934–14939 (2017)
70. Winfield, A., Vanderelst, D.: An architecture for ethical robots inspired by the simulation theory of cognition. Cogn. Syst. Res. **48**(1), 56–66 (2018)
71. Abbasimehr, H., Tarokh, M.J.: Trust prediction in online communities employing neurofuzzy approach. Appl. Artif. Intell. **29**(77), 733–753 (2015)
72. Bass, B.M.: From transactional to transformational leadership: learning to share the vision. Org. Dyn. **18**, 19–31 (1990)
73. Hersey, P., Blanchard, K.H.: Life cycle theory of leadership. Training Develop. J. **23**(5), 26–34 (1969)
74. Mattsson, S., Fast-Berglund, Å.: How to support intuition in complex assembly? Procedia CIRP **50**, 624–628 (2016)
75. Capraro, V., Everett, J.A.C., Earp, B.D.: Priming intuition disfavors instrumental harm but not impartial beneficence. J. Exp. Soc. Psychol. **83**, 142–149 (2019)
76. Barrafrem, K., Hausfeld, J.: Tracing risky decisions for oneself and others: the role of intuition and deliberation. J. Econ. Psychol. (2019, in press). https://doi.org/10.1016/j.joep.2019.05.004

The Privacy Barrier for Blockchain in Logistics: First Lessons from the Port of Hamburg

Niels Hackius(✉), Sven Reimers, and Wolfgang Kersten

Hamburg University of Technology (TUHH), Institute of Business Logistics and General Management, 21071 Hamburg, Germany
{niels.hackius,sven.reimers,logu}@tuhh.de

Abstract. Blockchain technology is associated with greatly beneficial applications for supply chain and logistics (SC&L), two of which are to trace goods across many actors, and to decentralize asset transfers without needing an intermediary. As a first use-case, actors from the Port of Hamburg are planning to implement blockchain to improve the sea freight container release by providing a common data platform for sea freight carriers, terminals, truck companies, and freight forwarders. Currently, releasing containers from the port's terminals to trucks requires proof of ownership for the recipient to take custody. In practice, this proof passes through many hands causing duplication of information flow and ownership evidence. We conducted workshops and short interviews with experts providing first-hand insight into the use-case. Using blockchain in the process provides improvements such as traceable proof of ownership. The technology also faces barriers, with privacy concerns as one of the most prominent obstacles. A decentralized system could lead to business networks and company information being disclosed through data triangulation. We argue that privacy is a vital design consideration that affects the use of blockchain in SC&L generally.

Keywords: Blockchain · Port of Hamburg · Privacy · Seafreight

1 Introduction

Blockchain is a technology expected to provide many opportunities for actors in supply chain and logistics (SC&L). Blockchain essentially is a distributed digital ledger of transactions that cannot be tampered with due to the use of cryptographic methods [27,32,34]. The use of blockchain for tracing assets in SC&L has been outlined in literature for various industries [3,6,8,26,35,36]. The decentralization makes it especially useful in multi-stakeholder environments with short-lived business relationships [12,26,36,37]. The release of sea freight containers at a ports' terminal is such an environment.

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 45–61, 2019.
https://doi.org/10.1007/978-3-030-29821-0_4

In this study, we look at the first steps of a deployment in the port of Hamburg where blockchain is intended to hold the custody status of a sea freight container during its transport from a vessel through the terminal onto a truck. Releasing containers from the port's terminals to trucks requires proof of ownership to take custody. In practice, this proof passes through many hands, via email forwards or printed copies, causing multiplication of information flow and allowing duplication of the proof of ownership. In the port of Hamburg, a blockchain based system could potentially improve the security and the efficiency of the release of sea freight containers. There are first examples of similar set-ups, such as the one by the company T-Mining in the port of Antwerp which combines the sea freight container release with digitalized documentation [31].

While several conceptual models for blockchain in SC&L are available, there is a lack of experience about implementation in practice. Prior to this study, we were able to identify three empirical, qualitative reports on blockchain use in marine SC&L [14,19,24]. These reports contribute to the question of blockchain use in SC&L and offer practical insights, however, they focus on particular conceptual models or technical tests. We investigate the managerial aspects of blockchain in SC&L, addressing opportunities and barriers. Filling this knowledge gap is a necessary precondition to achieve potential benefits of blockchain in a real-world SC&L context.

Qualitative data was collected from a deployment project currently in progress at the port of Hamburg. The data helped to identify possible opportunities for companies to profit from such a system even beyond the port environment. It also revealed extensive reservations, especially regarding transaction privacy. The results of this paper can guide future practical implementations, as well as a contribution of evidence towards theory building for blockchain in SC&L.

The remainder of the paper is structured as follows: After introducing the features of blockchain, we outline the qualitative survey method we used, and show how we augmented the results of the workshop by conducting short interviews with the stakeholders. Next, we present the results of this survey outlining the current and the planned process, as well as deriving specific and general opportunities and barriers that the interviewed experts noted. We conclude with a discussion of the findings, presenting a summary of practical problems implementers face in real life.

2 Theoretical Background

2.1 Basics of Blockchain Transactions

Blockchain is a distributed digital ledger of transactions that cannot be tampered with due to the use of cryptographic methods [27]. Outlined in Fig. 1, three basic properties of blockchain are relevant in management: (1) The data on a blockchain is stored in a decentralized manner and distributed to its members through a peer-to-peer network. No central authority has power over this data sharing; every member has its own, directly accessible copy. (2) Transactions are created by signing them and propagating them through the network. Public-private-key-cryptography is used to do this. Every participant has at least one public key with a corresponding private key. The public key serves as a public address and identifier, and is recorded with the transactions. Consequently, only the owner of the private key can initiate a transaction. In a business network, the owners of a public key are most likely known. (3) The transactions stored on a blockchain are immutable. A consensus-algorithm provides the members with a means of checking the validity of a transaction. The transactions are then stored in such a cryptographically interlinked way that it would be complicated to create fraudulent transactions or change past transactions.

Fig. 1. Basic properties of blockchain (based on [12])

There are many different software implementations of blockchain. They can be categorized in public (permissionless) and private (permissioned) blockchains. Public blockchains, like Bitcoin or Ethereum, are open to everyone who would like to join. In contrast, private blockchains feature a type of access control that restricts reading or writing, and even then, these access controls often require a central authority [27,32]. However, in consortia or business networks, the permissioned access is generally more interesting, because they often provide additional features the companies need.

In business, this kind of system can provide advantages by eliminating the need for a trusted intermediary party or a common data platform provider [22]. Asset transfers can be recorded without a trusted third party or an intermediar being required [34,38]. If every transaction involving goods were to be recorded on such a blockchain, this would help in tracking and tracing goods in SC&L and supporting anti-counterfeiting measures without the necessity of a central provider and the related onboard processes [8]. During the past few years these facilities that the technology promises have been observed very closely in SC&L. In this period blockchain has become a proper hype in the industry, yielding many new companies and projects [26].

2.2 Competitive Intelligence and Business Privacy

Competitive intelligence, a part of business intelligence, includes the collection, evaluation, and aggregation of information about markets and competitors [4,28]. A core task is gathering market insights about the competitor's strengths, weaknesses, products, services, and strategies [23]. Companies hope to achieve competitive advantages by gathering this knowledge or benchmarking their products [23].

Naturally, businesses seek to maintain privacy and to keep their activities secret from their competitors, particularly business processes or relationships [7]. However, cross-company data sharing is always susceptible to leaking this sensitive information. Especially in a network, like blockchain, where the functionality of verification and validation make transactions is decentralized and available to all participants [37]. While data security approaches can safeguard against unauthorized access even in this context, they cannot address the conclusions that can be drawn from the transaction data and metadata required for a functioning system [33,37].

3 Method

In this paper, we document an existing process and analyze barriers and opportunities of a possible transition to a blockchain based solution. We outline how the process is currently designed and what the shortcomings are. Also, specifically because the process involves various stakeholders from different companies, we identify the possible opportunities and barriers of blockchain.

In order to gain extended insight into these processes and build the related expert knowledge, a qualitative approach is appropriate [20,21,29]. In the context of this setup, we chose a two-fold way of gathering the data. In a first step we recorded the outcomes of four workshops, and in second step we conducted short interviews to record the experts' expectations and perceived risks.

The workshop outcomes were documented by creating reports. These reports were then used to derive the current and the expected process, as well as outlining the blockchain use-case. The workshops also helped the participants to gain a common understanding of the blockchain technology, the process handling by other stakeholders, as well as other stakeholders' challenges with the current and the planned process.

We conducted the interviews using a semi-structured interview guideline and we subsequently coded them as outlined in Mieg and Näf [21] and Mayring [20]. During the interviews, we asked the participants about the outcomes and profits they expected of the project, the problems they perceived with the blockchain implementation in that context, and their evaluation of the latter. In total, five interviews were conducted, each with a duration of between 11 and 43 min. A list of the experts is given in Table 1. The recorded material yielded 78 codings and revealed 6 topics to be discussed.

Overall the qualitative approach yielded a thorough understanding of the current processes. The workshop revealed multiple challenges, opportunities, and hidden actors. The interviews then allowed for the completion of this investigation by providing detailed insight, often focusing on possible risks associated with using blockchain technology.

Table 1. Sample of interviewed experts

#	Expert position[a]	Type of company
1	IT-coordination; process management	Container terminal operator
2	Internal IT solutions	Container terminal operator
3	Business intelligence analyst	Ocean freight carrier
4	IT consulting; IT project manager; IT project manager	Ocean freight carrier
5	Manager IT applications	Logistics service provider

[a](multiple expert of the same company separated by semicolon)

4 Results

Roughly 2.7 million containers were imported through the port of Hamburg in 2017 [2]. The containers are discharged from the sea freight vessels and moved to a terminal for temporary storage. The terminals' storage capacity is limited, therefore transport to their destination, in more than half of the cases by truck, has to proceed quickly.

We outlined the current and blockchain based processes of releasing containers from the terminal to a truck during four workshops, which lasted between two to three hours each. As shown in Table 2, seven parties were involved, which meant that multiple perspectives would be included. This first step, illustrated in Fig. 2, yielded the current process, and it disclosed five stakeholder groups, outlined in Table 3.

Table 2. Workshop participants in a container release process

Participant company	Workshops
Container terminal operator 1	Workshops 1–4
Container terminal operator 2	Workshops 1–4
Sea freight carrier 1	Workshop 1 and 3
Sea freight carrier 2	Workshops 1–4
Logistics service provider 1	Workshops 1–4
Logistics service provider 2	Workshop 1
Software company	Workshops 1–4
Ship broker	Workshop 1

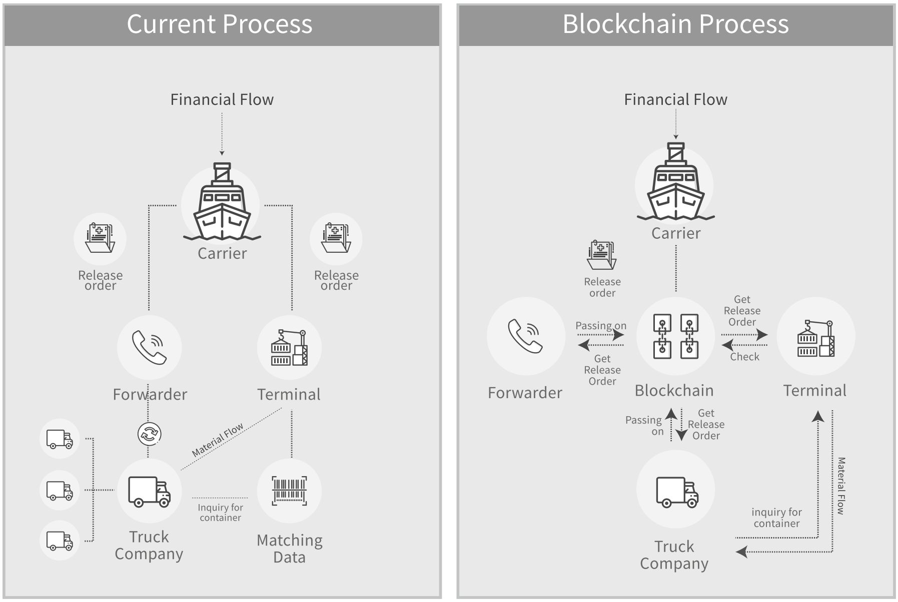


Fig. 2. Information and material flow in the current process and in a process using blockchain

Table 3. Stakeholders in a container release process

Roles	Explanation
Sea freight carrier	The sea freight carrier starts the release order process if the invoices have been paid. Next, the carrier authorizes the forwarder to pick up the container
Forwarder	The forwarder has the opportunity to hire a subcontractor to pick up the container at the terminal. In hiring a subcontractor, the forwarder assigns the collection right to the subcontractor
Truck company	The picking up of the container at the terminal is organized by the truck owner's company. This company assigns a truck driver to pick up the container at the terminal
Trucker	The truck driver has to identify himself at the terminal to pick up the container
Terminal	The terminal discharges the container form the ship, and has to ensure that the container is distributed to the right person, i.e. the driver with the right to collect a specific container. In order to check the truck driver's data the terminal receives the data from the carrier. The data sent by the carrier must match the data from the truck driver

The current process was found to be structured as follows: The company operating the container, i.e. the sea freight carrier, creates a digital "container release order" before the actual physical container is discharged. The container release order waives the requirement of presenting the actual bill of lading and facilitates a paperless pickup. This container release order is initially sent to only two parties, namely an official at the terminal of discharge that is to hold the container, and the forwarder, who in most cases is a logistics service provider (LSP) tasked with organizing the overland transport.

The digital container release order contains limited information about the container (see Table 4). The order's essential function is to associate the container number with the release reference number. The release reference number is important for taking custody of the container: the truck which the LSP tasks with transporting the container, has to present the release reference number at the terminal. The LSP cannot take custody of the container until the terminal has ascertained that the release reference number is the same as the LSP's.

LSPs typically commission a subcontractor to undertake the actual truck road haul. The landscape of companies providing these hauling services is very diverse since the service they provide is very standardized and highly competitive [5]. A large share of these services is sold on the spot market shortly before they are needed, and such service providers are easily replaced [5]. During daily operations, this implies a cascade of subcontractors who eventually get in touch with an actual truck operator. The release reference number has to propagate through this cascade for a particular truck operator to be able to pick up the container. By this time many parties have access to the release reference number that initially was known to only two parties.

Table 4. Data structure of a container release order

#	Field	Description
1	Container number	Worldwide unique number which defines every single container
2	Release reference number	Alpha-numeric index for uniquely defining every single container's release
3	Validity date	Date on which the release order expires
4	Carrier	A company that operates container transports
5	Terminal	A plant where containers are handled
6	Forwarder	An organization which dispatches goods

4.1 Opportunities of Blockchain

Unifying the Flow of Information. Lambert et al. outline three flows to be managed in every supply chain process: the material flow, the financial flow, and the information flow [17]. As there is an established financial flow regarding port and terminal handling fees, these fees are handled independently before the process starts. The material flow itself is simple in that the container is passed from the sea freight vessel to the terminal, and then loaded onto a truck.

The information flow, however, is not as straightforward. Both the freight forwarder and the terminal officials are provided with a copy of the same data, which means there is data duplication. The duplication makes it hard to propagate changes to the container release order, for example when changing validity or revoking the release order. Blockchain technology could enable the various stakeholders to work while using the same data as a basis. All stakeholders could have access to the same data and update it as the process progresses.

This would be beneficial to all stakeholders, because not only would their data basis be the same, but they would also achieve new communication possibilities. This is pointed out by one expert:

> We believe that we can access data more easily. Today the whole data transfer is very file based. [...] Through blockchain we can directly access the data with our programs. We do not have to take care of data transmission, because the data is replicated on the node and exactly the same as on all other nodes. Now that means we do not have to take care of access, the whole EDI stream that is needed today. We will not need any of that anymore.
>
> — #1, Container terminal operator

Maintaining Standardized Communication. That companies share the same data, is essential for an efficient supply chain network [18]. The Electronic Data Exchange (EDI) provides an electronic communication standard to achieve this data exchange, ensuring that the data is structured and not compromised by human error [10]. EDI is implemented through various standards, e.g., the UN/EDIFACT or GS1 EDI, as used in the logistics sector to automate interaction with partners [30]. The information flows analyzed and shown in Fig. 1 partially implement EDI. We found that the information flow to the terminals typically uses a technical interface which implements an EDIFACT standard. However, the information flow to the freight forwarders uses various channels, typically e-mail messages or phone calls. Only a few freight forwarders implement an EDIFACT based interface. In subsequent communication, from the freight forwarder to and between the multiple levels of subcontractors, we found that, typically, human-readable, non-standardized e-mail and phone messages are used.

The use of blockchain technology facilitates information flow in both directions. On the one hand, the communication between all participants can be standardized according to a communications standard. Fiaidhi et al. show that even an EDIFACT standard could be used [10]. On the other hand, an interface accessible to all kinds of stakeholders can be provided. Decentralized applications expose blockchain based data in a human-readable format through an internet browser or smartphone application, while the possibility of automation by other stakeholders later in the process is maintained.

Improving Formal Security. Space as well as terminal capacity is limited at the Hamburg port. In order to manage the steady incoming flow of trucks, since 2016, the pre-registration for a time-slot at the terminal through a system called “TruckGate” is mandatory for most Hamburg port terminals. TruckGate uses a proprietary EDI communication protocol called “TR02”, and it can be connected to through a human-readable web interface or through an API.

In order to register for a time-slot, the release reference number has to be provided, as well as a login ID or the trucker’s card ID. As outlined, the container release order information and the release reference number it contains are not personalized or encrypted. Every entity that has knowledge of the release reference number can initiate the container release at the terminal. We found that every subcontractor involved has the opportunity to release the container at any time.

This undesirable opportunity is somewhat mitigated by the fact that a formal registration with an attached know-your-customer process is required. Moreover, the terminal keeps photo records of each container pick-up. The terminal also matches the release reference number from the container release order with the number provided by the truck driver. However, currently this is a loose end as there is no way of telling for the terminal whether this truly is the up-to-date custodian of the container to be released. As the expert outlines:

> At the moment we get it [the container reference number] from the carrier. On the other side the forwarder passes it to a trucking company, and at some point, someone arrives at our door with the reference number. And then we compare: Is that the correct code? [...] then the truck drives through an OCR-gate[1], where it is photographed and – using text recognition – the container numbers and the license plate are recorded.
>
> — #2, Container terminal operator

Uncertainty about ownership can be reduced by using blockchain. The blockchain infrastructure would make it possible to assign the current ownership to only one specific public key at a time. The container would then only be released by the entity holding the private key to this address. Moreover, this assignment would provide additional security against unauthorized access by third parties. While an imposter could gain knowledge of the release reference number by eavesdropping on the email or phone communication of any company involved in subcontracting the container transport, in a blockchain solution deceptive access would only be possible if the private key of a single, specific entity were leaked.

4.2 Barriers to Implementing Blockchain

Privacy: Business Capacity and Business Network Disclosure. Gathering market insights is one of the core functions of competitive intelligence. Creating benchmarks, comparing offers, and identifying possible operative advantages are not limited to manufacturing companies. Although these techniques

[1] The optical character recognition (OCR) allows to convert the licence plate numbers and container numbers from image to text.

are typically used in market competition, blockchain now offers a common data platform that is useful to competitors, customers, and partners alike. The content of each transaction put on a blockchain can be encrypted. However, the major blockchain implementations currently available on the market use public key based pseudonyms to identify the party creating the transaction and the party this transaction is intended for [16].

In a simple setup of a blockchain implementation the complete transaction history is stored on a blockchain in plain text, facilitating easy data analysis. This analysis then reveals the public key addresses of the trading parties, the number of transactions, and possibly the content of the transaction. Even if the entities associated with the public key addresses are not known, their transaction types or patterns could lead to their identification, as illustrated in Fig. 3.

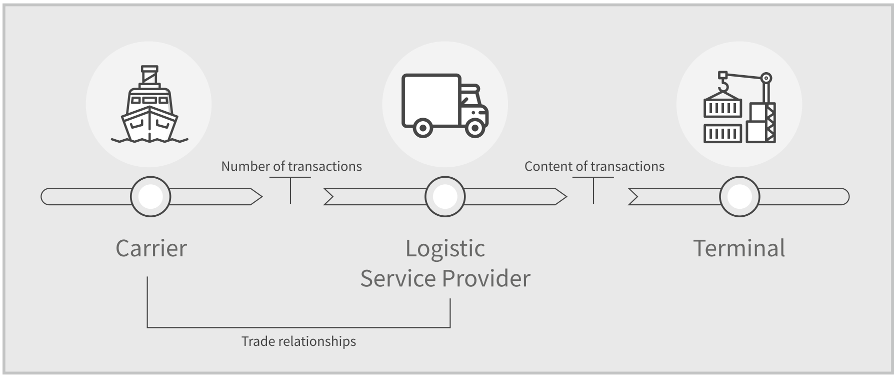


Fig. 3. Analysis of the blockchain data could allow for identification of trading parties and related volumes.

In the use-case we research, the participants expect that their public keys will soon be well known and connected to their company. The meta-data of the created transactions will consequently be available to their competitors for analysis. Moreover, a rigorous tracking of the transactions associated with the ownership of a container reference number will reveal parts of the freight forwarders' subcontractor network. Entity relationships and meta-data that otherwise remained unknown to the rest of the network, now become visible as single transactions. In the words of some stakeholders:

> ...in principle it is one of the biggest worries that everyone can access everything – and maybe use that information to see who is doing business with whom and to what extent. Is there any business proposal you could make to those business partners, could you undercut them?
>
> — #3, Ocean freight carrier

> The quantitative figures – it is not so much about them. Maybe you don't want to necessarily share them, so you don't indicate how full or empty you are, but it is not the big issue. We are relatively open about those. It is more about detailed shipment data, for example, who the recipient or the sender of the goods is.
>
> — #1, Container terminal operator

In combination with other public data sources (e.g. shipping trackers, customs information, container tracking), the number of transactions could give insight into the transshipment volumes of sea freight carriers, freight forwarders, and their subcontractors. Moreover, increased transparency through a digital transaction trail could reveal business relationships between companies. We found that the experts expect this knowledge to generate increased competition within the port of Hamburg subsystem.

Lack of Digtialization in Logistics Companies. A blockchain solution that connects all participants using different digital infrastructures, requires all participants to have and use digital access. Our interviews revealed this assumption of universal digital access to be false, as the interviewed experts claimed that many drivers either do not have access to the internet at all or that the use of a private key accessible only to them, would be too difficult for them to manage.

Kersten et al. show that digitalization in German logistics companies is somewhat limited [15]. Otherwise, the majority of Germans own a smartphone and use it a lot [1]. Most likely this is true for truck drivers as well. It seems likely, the companies that employ the truck drivers are unwilling to pay for an infrastructure that the driver could also use for private purposes. Furthermore, because a permissioned or private blockchain solution is typically preferred, a certain on-boarding and know-your-customer process would be required.

A blockchain implementation that considers this would, in this case, end at the TruckGate. While this does not cover the full end-to-end processes, implementing a blockchain has some potential to reduce attack vectors and would add the possibility for the TruckGate platform to check whether the container release number for which the slot is requested, actually belongs to that company.

Lack of Blockchain Technology Readiness. A third barrier that was revealed relates to the fear of technology immaturity. While new technologies are constantly deployed, blockchain is still to some extent considered to be an immature technology concept which is split up in many implementations that evolve very fast [11]. A part of this discussion concerns the transaction speed and the speed of the consensus algorithm – the existing blockchain implementations are assumed to be rather slow. During a normal day at the port, one terminal operator handles approximately 2000 trucks per day, and up to 250 trucks per hour [25]. We have limited knowledge of the actual figures across all terminals of the port, as well as of the depth of the subcontractor chains. Importantly, the container reference number has to be exchangeable until very shortly before the actual container-pick-up. Hence, the experts are concerned that a deployed system might be fast enough. As one expert points out:

> So, we think, one thing that is relevant is the speed. [...] In real life the truck driver that picks up the container has to be switched on short notice. So, the technology has to be quick enough for truck drivers to get the container that they are supposed to get immediately – and not maybe an hour later.
>
> — #5, Logistics Service provider

Although the blockchain technology is still considered to be immature, it is likely that these transaction rates are achievable. It is possible that the time it takes to build a block and communicate it to the user front ends is longer than the communication times in the current setup of phone calls and short messages. However, it seems unlikely that it will hamper the business process.

5 Discussion

Our study has revealed barriers and opportunities based on an actual blockchain project in the process of being implemented. The properties of blockchain lead to certain advantages: We found that decentralized architecture of blockchain will enable each stakeholder to access to the same dataset. Effectively this will allow the cross-company tracing of container custody and provide a common platform unifying the information flow. The entries in this database are verified and immutable which will allow the blockchain participants to be sure that only one proof of ownership exists. In the future, this system could be used end-to-end allowing the truck driver to prove ownership of the container through blockchain themselves.

On the other hand, there are disadvantages: blockchain comes with challenges for the companies which eventually prevent the introduction of this mutual data storage. According to the experts, the technology itself presents two barriers to the practitioners: (1) a possible lack of privacy and hence the disclosure of business relationships, (2) immaturity of the blockchain implementation itself, yielding possible problems with performance. A third barrier is the technology readiness of the companies themselves. Many of the small companies lack the resources or the knowledge to implement software on their own.

A final working solution for this process will have to address all three barriers to leverage the potential of blockchain in this process. A possible alternative would be a centralized platform provided by an IT company. In Table 5 we summarize the differences between these approaches. It can be concluded that the main differences are in the extent to which the involved parties can trust the platform provider, and how they can additionally profit from the blockchain system itself. The experts note this saying:

Table 5. Comparison between the current solution, a blockchain solution and a platform solution

	Current solution	Blockchain	Platform solution
Data			
Storage	All participants	All partcipants	Platform provider
Duplication	Disconnected datasets that can be copied	One connected dataset that is copied many times	No duplication
Access	Anytime, if the information is available	All partcipants, any time	Any time, access restrictions at provider's discretion
Communication Standards	EDI, E-mail, Telephone, Fax	Digital only, API depends on the implementation	Depends on the platform provider
Ownership			
Verifying current owner	not possible	possible	possible
Tracking past owners	not possible	possible	possible
Digitalization			
Need for Internet access	not needed	required	required
Digitalizing related business processes	not needed	required	required
Privacy			
Disclosure of Business relationships	Difficult to trace	Can be encrypted, depending on the implementation; triangulation might be possible	Access restrictions, platform provider has full access
Container reference number	Can be shared	not needed	Access restrictions, platform provider has full access
Transshipment volume	Difficult, manual tracing	Triangulation might be possible	Access restrictions, platform provider has full access

> ...the only thing that justifies blockchain is the real decentrality ...
>
> — #2, Container terminal operator

> ...the benefit is if you really extend it to the entire supply chain later and make processes leaner
>
> — #3, Ocean freight carrier

Providing privacy by shielding business information against retrieval by data triangulation stands out as a challenge for blockchain technology. Currently, to the best of our knowledge, it is not possible to achieve this kind of privacy. This is partially rooted in the way current blockchain implementations achieve consensus and immutability. There are promising approaches in public and permissioned/private blockchain space. Ethereum has at least two initiatives, Plasma or Raiden, which try to address this by implementing non-interactive zero-knowledge proofs (zk-snarks). The development projects Hawk [16] and Enigma, try to provide an entirely private approach [13]. Quorum, an implementation targeting the permissioned/private space, also claims to work on implementing zk-snarks [9]. IBM driven Hyperledger fabric offers a service called "identity mixer". However, all of these plans are either still under development, or require a centralized trusted entity.

We found that, for companies, at the moment privacy concerns outweigh the advantages blockchain could have for them. Moreover, the lack of digitalization in logistics companies, as well as the immaturity of existing blockchain implementations, can be a barrier to the use of blockchain in practice.

6 Limitations and Future Research

Our study holds some limitations, which should be taken into consideration when evaluating the findings. First, the studied use-case is still in progress. It is possible that during the pilot and the implementation phases of the project barriers we discovered can be mitigated, or that expected advantages are not realized. Second, the interview sample is limited to the project participants and therefore provides insight into the early learning of logistics experts currently involved in the project. It is likely that experts with more blockchain experience would have noted further aspects of advantages or limitations. We plan to address this in the future by piloting first solutions with the experts. Furthermore, experts beyond the project are to be involved in order to mitigate biases.

Research in the blockchain sector is evolving very rapidly, thus the maturity of blockchain implementations is likely to improve. However, the existing research is still very limited and focused on very particular topics. In order to create improved blockchain implementations, we need more empirical reports that outline the requirements and problems of the implementing practitioners.

7 Conclusion

In this paper, we took a closer look at a first SC&L use-case regarding container release on a blockchain in the port of Hamburg. Doing so, we intended to report initial lessons and possible implications of implementing a blockchain based solution in SC&L. We presented two advantages and three barriers as findings from our workshops and short interviews. We found that the experts consider the unification of the information flow a possible advantage as it could make the process more efficient, transparent, and less redundant. However, we also found that increased transparency and the required digital proficiency, as well as the possible disclosure of partners and subcontractors, present a major barrier for companies. We compared the current solution to the blockchain solution and a central platform approach. A central platform solution approach currently is much easier to implement while delivering similar results, but has implications regarding bargaining power.

Our contribution will help researchers, implementers, and developers to design future blockchain implementations with the implications of business intelligence and limited digitalization in mind. Even though the knowledge we gain from practice is still somewhat limited, it is advisable for companies to engage in case-studies and field testing to see whether blockchain can benefit their processes and information flows. Overall, the blockchain technology holds opportunities regarding process optimization, with the possibility of being used across the whole supply chain.

Funding Disclosure. This paper was funded as part of the "Release Order based on Blockchain" (ROboB)-project. ROboB is a joint project within the framework of the IHATEC program, which is funded by the Federal Ministry of Transport and Digital Infrastructure (BMVI).

References

1. Bitkom Research, Smartphonemarkt: Konjunktur und Trends (2018). https://bit.ly/Bitkom-Pressekonferenz-Smartphone-Markt-22-02-2018
2. Port of Hamburg, Container handling 1990 to 2017. Port of Hamburg, Hamburg (2019). https://www.hafen-hamburg.de/en/statistics/containerhandling
3. Agrawal, T.K., Sharma, A., Kumar, V.: Blockchain-based secured traceability system for textile and clothing supply chain. In: Thomassey, S., Zeng, X. (eds.) Artificial Intelligence for Fashion Industry in the Big Data Era, vol. 05, pp. 197–208. Springer, Singapore (2018). https://doi.org/10.1007/978-981-13-0080-6_10
4. Baars, H., Kemper, H.G.: Management support with structured and unstructured data—An integrated business intelligence framework. Inf. Syst. Manage. **25**(2), 132–148. https://doi.org/10.1080/10580530801941058
5. Borgström, B., Jensen, L.M., Hertz, S.: Road haulier competition: Implications for supply chain integration. In: Gammelgaard, B., Prockl, G., Kinra, A., Aastrup, J., Holm, A.P., Schramm, H.-J., Hsuan, J., Malouf, M., Wieland, A. (eds) 26th Conference of the Nordic Logistics Research Network, Copenhagen (2014)
6. Cartier, L.E., Ali, S.H., Krzemnicki, M.S.: Blockchain, chain of custody and trace elements: An overview of tracking and traceability opportunities in the gem industry. J. Gemmol. **36**(3), 212–227 (2018). https://doi.org/10.15506/JoG.2018.36.3.212
7. Camenisch, J., Sommer, D., Leenes, R.: Digital Privacy: PRIME - Privacy and Idendity Management for Europe. Springer, Heidelberg, Dordrecht, London, New York (2011)
8. Casino, F., Dasaklis, T.K., Patsakis, C.: A systematic literature review of blockchain-based applications: Current status, classification and openissues. Telematics Inform. **36**, 55–81 (2019). https://doi.org/10.1016/j.tele.2018.11.006
9. Espel, T., Guillaume, R., Laurent, K.: Proposal for Protocol on a Quorum blockchain with Zero Knowledge, Cryptology ePrint Archive (2017). https://eprint.iacr.org/2017/1093.pdf

10. Fiaidhi, J., Mohammed, S., Mohammed, S.: EDI with blockchain as an enabler for extreme automation. IT Prof. **20**(4), 66–72 (2018). https://doi.org/10.1109/MITP.2018.043141671
11. Fridgen, G., Lockl, J., Radszuwill, S., Rieger, A., Schweizer, A., Urbach, N.: A solution in search of a problem: A method for the development of blockchain use. In: Americas Conference on Information Systems, New Orleans, LA (2018)
12. Hackius, N., Petersen, M.: Blockchain in logistics and supply chain: Trick or treat? In: Kersten, W., Blecker, T., Ringle, C.M. (eds) Digitalization in Supply Chain Management and Logistics, epubli, Hamburg, vol. 23 (2017). https://doi.org/10.15480/882.1444
13. Hepp, T., Sharinghousen, M., Ehret, P., Schoenhals, A., Gipp, B.: On-chain vs. off-chain storage for supply- and blockchain integration. IT-Inf. Technol. **60**(5–6), 283–291 (2018). https://doi.org/10.1515/itit-2018-0019
14. Jabbar, K., Bjørn, P.: Infrastructural grind. In: GROUP 2018, ACM Association for Computing Machinery, Part F139583, Sanibel Island, FL, pp. 297–308 (2018). https://doi.org/10.1145/3148330.3148345
15. Kersten, W., Seiter, M., von See, B., Hackius, N., Maurer, T.: Trends and Strategies in Logistics and Supply Chain Management – Digital Transformation Opportunities, DVZ, Bremen (2017). https://logistiktrends.bvl.de
16. Kosba, A., Miller, A., Shi, E., Wen, Z., Papamanthou, C.: Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In: 2016 IEEE Symposium on Security and Privacy, pp. 839–858. IEEE, San Jose, CA (2016). https://doi.org/10.1109/SP.2016.55
17. Lambert, D.M., Cooper, M.C., Pagh, J.D.: Supply chain management: Implementation issues and research opportunities. Int. J. Logistics Manage. **9**(2), 1–20 (1998). https://doi.org/10.1108/09574099810805807
18. Lee, H.L., Whang, S.: Information sharing in a supply chain. Int. J. Technol. Manag. **1**(3–4), 79–93 (2000). https://doi.org/10.1504/IJMTM.2000.001329
19. Loklindt, C., Moeller, M.P., Kinra, A.: How blockchain could be implemented for exchanging documentation in the shipping industry. In: Freitag, M., Kotzab, H., Pannek, J. (eds.) Dynamics in Logistics, pp 194–198. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-74225-0_27
20. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken, 12th edn. Beltz Pädagogik, Beltz, Weinheim (2015)
21. Mieg, H.A., Näf, M.: Experteninterviews, 2 Aufl., Institut für Mensch-Umwelt-Systeme (HES), ETH Zürich, Zürich (2005)
22. Morabito, V.: Business Innovation Through Blockchain. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-48478-5
23. Müller, R.M., Lenz, H.J.: Business Intelligence. Springer, Heidelberg (2013)
24. Nærland, K., Beck, R., Müller-Bloch, C., Palmund, S.: Blockchain to rule the waves - nascent design principles for reducing risk and uncertainty in decentralized environments. In: Kim, Y.J., Agarwal, R., Lee, J.K. (eds.) Proceedings of the International Conference on Information System: HCI, Association for Information Systems, Seoul (2018)
25. Naumann, J.P.: HHLA optimiert Lkw-durchlauf. DVZ, Hamburg (02 March, 2015). https://www.dvz.de/rubriken/see/detail/news/hhla-optimiert-lkw-durchlauf.html
26. Petersen, M., Hackius, N., von See, B.: Mapping the sea of opportunities: Blockchain in supply chain and logistics. IT-Inform. Technol. **60**(5–6), 263–271 (2018). https://doi.org/10.1515/itit-2017-0031

27. Pilkington, M.: Blockchain technology: Principles and applications. In: Olleros, F.X., Zhegu, M. (eds.) Research Handbook on Digital Transformations, pp. 1–39. Edward Elgar Publishing, Cheltenham (2016)
28. Porter, M.E.: Competitive strategy: Techniques for analyzing industries and competitors. Free Press, New York, NY (1998)
29. Punch, K.F.: Introduction to Social Research: Quantitative and Qualitative Approaches, 3rd edn. Sage, Los Angeles, London, Washington DC (2014)
30. Reichert, P.: Comarch EDI platform case study: The advanced electronic datainterchange hub as a supply-chain performance booster. In: Golinska, P. (ed.) Logistics Operations, Supply Chain Management and Sustainability, pp. 143–155. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07287-6_10
31. Stahlbock, R., Heilig, L., Voß, S.: Blockchain in der maritimen Logistik. HMD Praxis der Wirtschaftsinformatik **55**(6), 1185–1203 (2018). https://doi.org/10.1365/s40702-018-00464-8
32. Swan, M.: Blockchain: Blueprint for a New Economy. O'Reilly, Beijing (2015)
33. Sweeney, L.: k-anonymity: A model for protecting privacy. Int. J. Uncertainty, Fuzziness and Knowl. based Syst. **10**(5), 557–570 (2002). https://doi.org/10.1142/S0218488502001648
34. Tapscott, D., Tapscott, A.: Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World. Penguin, New York, NY (2016)
35. Tian, F.: A supply chain traceability system for food safety based on haccp, blockchain & internet of things. In: Tang, J., Chen, J., Cai, X. (eds.) 2017 14TH International Conference on Service Systems and Service Management (ICSSSM), IEEE, New York, NY (2017). https://doi.org/10.1109/ICSSSM.2017.7996119
36. Wang, Y., Han, J.H., Beynon-Davies, P.: Understanding block chain technology for future supply chains: A systematic literature review and research agenda. Supply Chain Manage. Int. J. **1**(103), 2 (2018). https://doi.org/10.1108/SCM-03-2018-0148,
37. Wust, K., Gervais, A.: Do you need a blockchain? In: 2018 Crypto ValleyConference on Blockchain Technology (CVCBT), pp. 45–54. IEEE, Zug (2018). https://doi.org/10.1109/CVCBT.2018.00011
38. Zyskind, G., Nathan, O., Pentland, A.S.: Decentralizing privacy: Using blockchain to protect personal data. In: 2015 IEEE Security and Privacy Workshops, pp. 180–184, IEEE, San Jose, CA (2015). https://doi.org/10.1109/SPW.2015.27

A Literature Review on the State of the Art of Multi-agent Systems in Supply Chain Management

Alexandra Fiedler[1(✉)], Dirk Sackmann[1(✉)], and Hans-Dietrich Haasis[2(✉)]

[1] Hochschule Merseburg, Eberhard-Leibnitz-Str. 2, 06217 Merseburg, Germany
[2] Universität Bremen, Bibliothekstraße 1, 28359 Bremen, Germany

Abstract. Agent-based software attracts great interest in industry and research, the main reasons being the efficiency, robustness and complexity minimization of such multi-agent systems (MASs). In addition, the application possibilities are varied. This paper presents an overview of the different areas and topics in which MASs are used and specifically addresses the question of how MASs are used in supply chain management (SCM). For this purpose, the identified studies are classified in the supply chain planning matrix and gaps in research are subsequently identified.

Keywords: Multi-agent systems · Supply chain management · Literature review

1 Introduction

In an increasingly competitive industrial environment, companies must either diversify their product portfolio to expand their offer in the market or specialize in a business to gain efficiency. They are therefore increasingly in contact with a large number of partners, as the entire production and distribution of the products cannot be conducted by a single company. Companies thus form a network, commonly referred to as a "supply chain". The management of this chain is realized through an exchange of information and the redistribution of activities between the various links that constitute it. SCM has been an important engine of corporate competitiveness for over twenty years. Initially, this concept was merely an extension of logistics practice to a larger number of partners before or after the operation of a company (Bechtel and Jayaram 1997). Industry and research have taken initiatives to make decision makers aware of the savings that can be achieved through the implementation of cooperative relationships. Often there are organizational barriers between the partners of such a supply chain, and information flows can be limited to the extent that full centralized control of the material flows in a supply chain may not be possible or desirable. As a result, most companies use decentralized control to manage the various assets in a supply chain (Lee and Billington 1993). Distributed supply chains (DSCs) can be defined as networks of autonomous components that operate in a competitive or collaborative environment in which no hierarchy of decision making is enforced, and initiatives are taken by each partner to achieve a common goal (Ghirardi et al. 2008). Planning a DSC

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 62–74, 2019.
https://doi.org/10.1007/978-3-030-29821-0_5

is a complex process that involves multiple constraints, including collaboration between the various organizational units of the system. The type of product to be transported, the category of targeted customers and the target group must be taken into account.

Since conventional approaches are generally monolithic and their concept is centralized, the current applications are based on multi-agent systems. MASs and standalone agents provide a new way to analyses, design and implement diverse applications, as they are part of distributed artificial intelligence and benefit from other scientific disciplines, such as cognitive sociology and social psychology. An agent in this context is a self-contained computer system that operates in a particular environment. It is able to act flexibly and autonomously to achieve its stated goals (Franklin and Graesser 1996). A population of agents who either collaborate or work against each other is referred to as a multi-agent system. It is also possible for both forms to occur in one system (Wooldridge 2009). Multiagent systems can be found in all application areas where characteristics such as situational awareness, openness, local control and local interaction are important (Klügl 2012):

- Systems in which software or hardware entities exist in an environment and interact with this and other entities;
- Systems in which the individual entities pursue their own goals and new entities can be added from outside;
- Systems in which a distribution of controls and data is inherent, or it makes sense to introduce such a distribution because the complexity can no longer be handled by a central control;
- Systems in which parallel, asynchronous processes are active.

The following study gives an overview of the variety of possible applications of MAS and describes in a second step the use of MAS in SCM.

2 Systematic Literature Research

To obtain an overview of the current state of research on MASs, a systematic literature search (SLR) was carried out in several stages. This is an established research methodology that helps to synthesize the existing knowledge in a systematic and transparent way and to support the identification of research gaps and the formulation of innovative research questions (Cooper and Valentine 2008). The first stage of the SLR was to gain an overview of the many forms and applications of MASs in research. The identified work would also be categorized after analysis. The second stage consisted of research with a focus on the field of production and logistics. The International Journal of Production Research (IJPR) generated the most results in this thematic orientation in the level 1 research. Therefore, the level 2 research, with the keywords "production, logistics and MAS", focused on this journal. After the categorization of these contributions, the articles of the SCM category in particular were examined and classified in the supply chain planning matrix.

3 Results

The results of the literature research are presented below. The procedure for the research is explained for both stages, and the subsequent analysis of the identified work is presented.

3.1 Stage 1

The aim of the SLR in level 1 was to obtain an overview of the different forms and research areas related to the subject of MASs. Since the research activities and thus the number of publications in this area were very numerous, the research focused on papers that contain corresponding literature analyses. Table 1 below shows the procedure underlying the SLR and the results. The difference between the total number of hits and the works classified as relevant is due to two issues. On the one hand, the result lists contained duplicate hits, that is, works that appeared in both databases. On the other hand, the search strategy produced hits that did not match the content of the MAS topic.

Table 1. Procedure in Level 1 of the SLR

Terms of the literature research			
Multi-agent systems	Literature review		
	Survey		
Search strategy	Database	Hit	Relevant
Multi-agent systems AND literature review	Business Source Premier	21	20
Multi-agent systems AND survey		43	41
Title, abstract, keywords: Multi-agent systems AND literature review	ScienceDirect	37	14
Title, abstract, keywords: Multi-agent systems AND survey		102	58
	Total:	203	133

These contributions were excluded from further analysis. For further consideration, the works were classified according to their thematic priorities. Here, the content of the debate included the following categories: information technology, production and logistics, economics, energy, medicine, networks, geography, behavioral science, transport, construction and architecture, and biology. Not every work could clearly be assigned to only one category. Some works spanned several thematic areas and were assigned to several categories. The graph in Fig. 1 shows that the majority of the articles reviewed considered the subject of MASs from the point of view of computer science. The topic has also received considerable attention in the area of production and logistics.

In the field of information technology (IT), MASs have been examined in various fields. The studies were assigned to the following areas as part of the analysis:

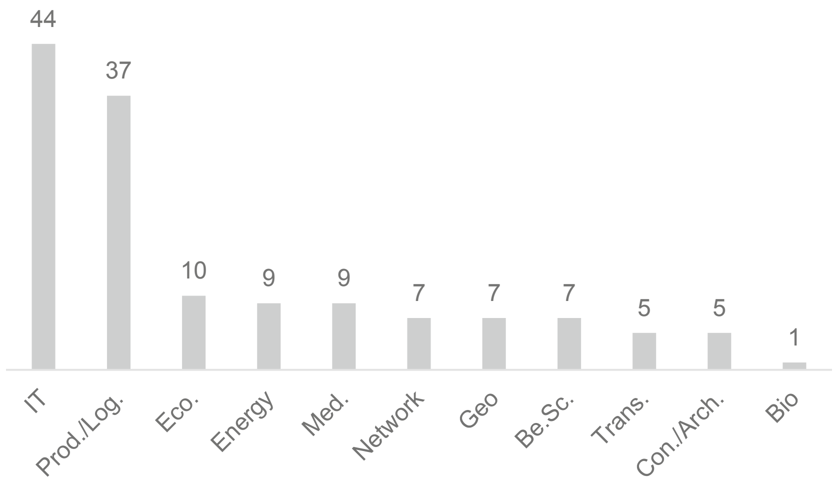


Fig. 1. Distribution of the studies on different topics

programming, artificial intelligence, simulation, robotics, agent self-assignment, data mining and knowledge management, security and complexity.

Work dealing with aspects of urban transport and transport policy is summarized in the transport (Trans.) category. The networks category covers topics such as the relationship between MASs and social networks, network management or even wireless sensor networks. Topics such as energy systems and safety, the state of charge of batteries and voltage analysis of power grids or smart grids are assigned to the category energy. Studies on the interaction between population and environment and studies on waste policy and urban development are the topics of the economics (Eco.) category. The construction and architecture (Con./Arch.) category deals with the application of MASs in building operations and the selection of contractors to provide a platform for the economic and environmental benefits of building renovations and the study of the properties of water supply systems.

MASs are also widely used in medicine (Med.), for example for cancer screening, as diabetes simulation software, for task planning in health organization or for the modelling of tuberculosis. The geography (Geo.) category deals with the simulation of human–environment interactions and the localization of high-volume zones in gas and oil reservoirs as well as soil analyses and land use change problems. The work in biology (Bio.) deals with models for huge amounts of experimental data. For example, MASs are used in the behavioral science (Be.Sc.) field to analyze religious extremism, to study the acceptance of language assistants, to accept water reuse or to analyze elections and voter behavior.

The contributions just presented in stage 1 of the literature analysis are distributed among 89 different journals and conference papers, of which 15 publication organs appear on the overall list of professional journals in VHB-JOURQUAL3. Table 2 shows these journals with the corresponding rating and the number of identified contributions.

Table 2. List of journals with JQ3 rating

Journal	Rating	Hits
IIE Transactions	A	1
Journal of Operations Management	A	1
ACM Computing Surveys	B	3
Annals of Operations Research	B	1
Artificial Intelligence	B	1
Decision Support Systems	B	1
Energy Policy	B	1
International Journal of Production Economics	B	3
International Journal of Production Research	B	12
Journal of Cleaner Production	B	1
Computers in Industry	C	2
Electronic Commerce Research and Applications	C	1
IEEE Transactions on Systems, Man, and Cybernetics	C	3
Journal of Enterprise Information Management	C	1
Computer Networks	k.R.	1

Most contributions are from the International Journal of Production Research (IJPR). All 12 of these articles could be assigned to the category production and logistics. In stage 2, the search for MASs in production and logistics should be deepened due to the high number of hits and the professional orientation of the authors. Because of the relatively high number of hits of the still general search of the previous stage, the analysis focused on the IJPR.

3.2 Stage 2

The search in the IJPR with the search words "multi-agent system" yielded 142 hits at the time of the request, of which 133 were classified as relevant. The contributions found were also categorized to identify the main topics based on the number of contributions. The categories that could be derived from the content of the contributions are shown in Fig. 2.

In the following, only the work of the supply chain management (SCM) category will be explained in detail to make the scope manageable. The 30 identified works in the category were analyzed on the one hand according to SCM aspects and on the other hand a consideration was made under MAS relevant aspects.

3.3 MAS Classification

When multiple agents work together, some additional tasks must be processed that are not directly productive but serve to improve the form of this collaboration (Ferber und Kirn 2001). From the perspective of MASs, coordination is a process in which agents become active to ensure that a community of individual agents acts coherently

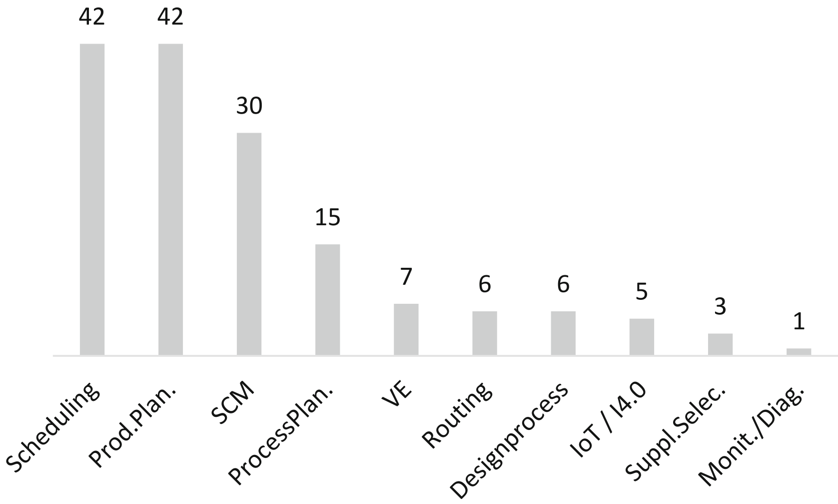


Fig. 2. Distribution of contributions to MAS in the IJPR

(Nwana et al. 1996). Coordination techniques are classified according to the three dimensions control, structure and attitude (Davidsson et al. 2005).

In the dimension of control, the authority relationships between the agents are recorded, which are either centralized or decentralized. The MAS structure corresponds to the set of agents that make up the MAS, their roles and the communication paths between the agents. The structure is either predefined, i.e. static (the number of agents or their roles do not change during execution) or dynamic. Finally, the dimension of agent attitude captures the behavior of agents that is either classified as benevolent (cooperative), i.e., that social laws and global goals are adhered to, or selfish (competitive), when the individual goals of agents, e.g., in a market economy, determine their behavior.

Before classification according to MAS can take place, it must be noted that of the 30 previously identified contributions, only 16 have developed, applied or modified a MAS. The remaining 14 contributions do not fit into the proposed classification, as they are either literature reviews or papers that deal primarily with SC aspects and only mention MAS as a possible technique for problem solving or decision support without using it. The results of the analysis are shown in Fig. 3.

The majority of projects therefore rely on decentralized control. Dai et al. (2014) use a Distributed Multiagent System (DMAS) for the simulation of supply chains. They propose a methodological framework for supply chain modeling based on fractal thinking and multi-agent technology for distributed simulation. Chen and Nof (2010) in turn use MAS to investigate the performance of a decentralized conflict and error detection and prediction model (CEDPM) in different networks. For this purpose, the authors develop a conflict and error detection and prediction protocol (CEDPP). The CEDPP is an agent-based protocol that defines how the CEDPM components communicate with each other.

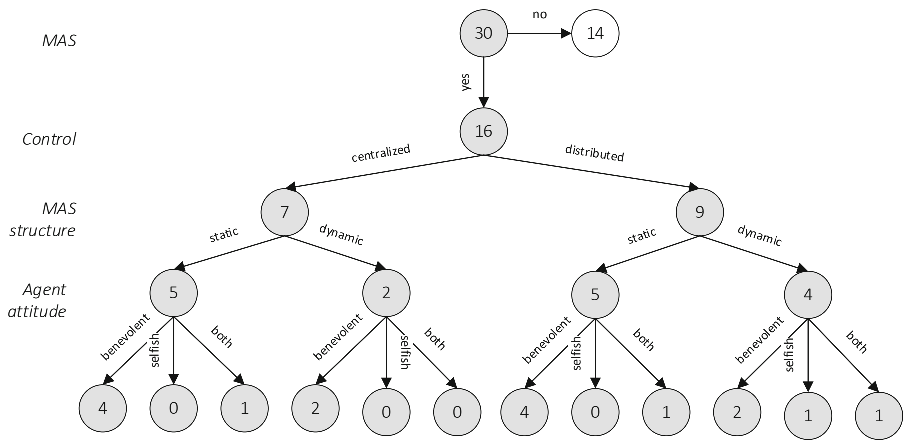


Fig. 3. Number of approaches to MAS classification

In addition, only 37.5% of applications make use of dynamic MAS structures, which is an often-cited strength of agent technology. Regarding agents, cooperative behaviors predominate. The MAS developed by Caridi et al. (2005) for the collaborative planning forecasting and replenishment (CPFR) process includes so-called intelligent agents. In the "learning" model, the agents are more intelligent than in the previous models because they can learn from the past. In particular, the learning model was developed to manage a product life cycle. The model is able to capture the moment when one product is mature while another is phased in. In contrast, Chan and Chan (2009) use a static MAS structure. In their agent-based model, they introduce delivery volume and maturity flexibility to reduce costs and improve fill rate in a distributed supply chain exposed to uncertain demand and supply.

In addition to the coordination techniques, an analysis was made of the maturity of the MAS, i.e. how complete and validated an application is. Parunak (2000) and further developed by Davidsson et al. (2005) suggest the following levels of maturity: the lowest maturity level is the conceptual proposal, the next level are simulation experiments, the field test level indicates that the test was performed in the environment in which the application is to be used and the last level, the deployed system, indicates that the system has been implemented in the real world and was or is in use.

Figure 4 visualizes the distribution showing that most MAS are in the phase of being tested in a simulation environment.

The three conceptual works are by Wang et al. (2011), Mishra et al. (2012) and Huang et al. (2006). Wang et al. (2011), for example, propose an approach to ontology-based knowledge representation with the goal of providing a semantic interoperable environment for the realization of automated negotiations in virtual enterprises. The proposed concept can be implemented by a MAS. The authors develop two agent-interaction protocols to retrieve semantic correspondence and negotiate offers.

The development of Blos et al. (2018) has already reached maturity to carry out a field experiment. The case study is based on an airline that transports goods from China

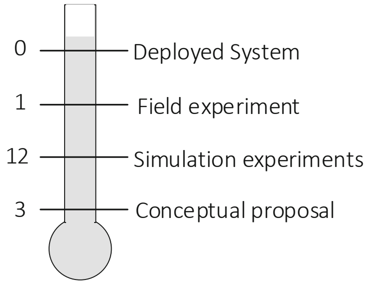


Fig. 4. Number of approaches to maturity level

to Brasil. Based on this data, two scenarios were created to test the proposed framework for designing supply chain disruptions management.

The twelve contributions (Dai et al. (2014), Yu et al. (2017a), Shukla and Kiridena (2016), Long (2014), Chan and Chan (2010), Li and Sheng (2011), Caridi et al. (2005), Confessore et al. (2008), Lu et al. (2005), Enjalbert et al. (2011), Chen and Nof (2010), Allwood and Lee (2005)) in which simulations were carried out are explained in more detail in the following chapter.

3.4 SCM Classification

The 30 contribution were now integrated into the supply chain planning matrix according to Fleischmann et al. (2005) and Rohde et al. (2000). The supply chain planning matrix classifies the planning tasks into the two dimensions “planning horizon” and “supply chain process”. Figure 5 shows the typical tasks that occur in most types of supply chains and their correspondence in the planning horizon.

As the figure shows, most work dealt with issues of long-term SCM. In particular, considerable attention has been paid to the cooperation of the partners involved and the choice of suppliers, as explained below by way of example. Huang et al. (2003), for instance, reviewed the research on the impact of sharing production information on the dynamics of the supply chain. The purpose of the review was to highlight the implications for the design and management of the supply chain, to identify the current research frontier and to examine the future trend in this area. Caridi et al. (2005) again proposed two multi-agent models and tested them for use in the collaborative planning forecasting and replenishment (CPFR) process. Lu et al. (2005) proposed an agent-based collaborative production framework to ensure the flexibility and real-time responsiveness of a supply chain system. Jain et al. (2009) investigated how various modelling techniques, such as agent technology, Petri nets, fuzzy logic and data mining, can be used to support dynamic supply chain configurations concerning vendor-related issues.

Chan and Chan (2010) reviewed the SCM work related to the supply chain dynamics of supply chain coordination. The focus was on inventory management issues. Yu et al. (2017a) presented a negotiation protocol specifically for the problem of selecting multi-product suppliers. The negotiation protocol is a hybrid multi-agent protocol of the combinatorial procurement auction protocol and the multi-bilateral

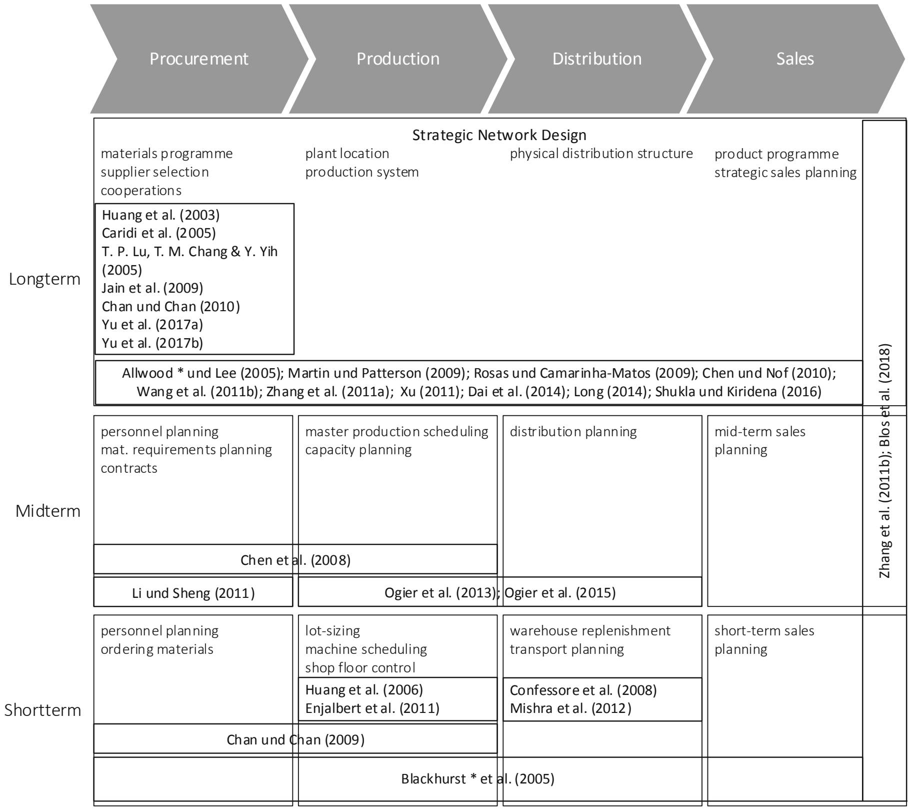


Fig. 5. Classification of the studies in the supply chain planning matrix according to Fleischmann et al. (2005)

negotiation protocol. Yu et al. (2017b) presented an MAS to show how different levels of collaboration affect service levels and how cloud computing can foster these levels of collaboration.

A range of work has dealt with long-term planning with the SC network design. For example, Allwood and Lee (2005) suggested an agent to investigate supply chain network dynamics' competitiveness. Martin and Patterson (2009) examined the use of common metrics to determine which are the most useful for measuring performance when companies implement SCM practices. The authors drove the search for suitable measures for the connection between the supply chain and the financial performance of a company. Rosas and Camarinha-Matos (2009) presented a modelling approach to assess a business's readiness to join a collaborative network. The goal of the research by Chen and Nof (2010) was to examine the performance of a newly developed decentralized conflict error detection and prediction model (CEDPM) in various networks. Wang et al. (2011) dealt with negotiating knowledge and introduced an ontology-based approach to knowledge representation to create a semantic interoperable environment for the realization of automatic negotiations in a virtual enterprise (VE). Xu (2011) and

Zhang et al. (2011a) considered the quality aspects of SCs. For this purpose, work on quality coordination in the supply chain, technology application, risk management in the supply chain and reliability control were examined. Dai et al. (2014) and Long (2014) were concerned with supply chain modelling and simulation: Dai et al. on the basis of a fractal perspective and Long through the integration of agent-based distributed simulation and an improved supply chain operations reference (SCOR) model. Likewise, Shukla and Kiridena (2016) dealt with the configuration of supply chain networks and proposed an appropriate framework.

Only four papers dealt with MASs for mid-term SC planning problems. Chen et al. (2008) developed a resource planning model with two profit-centered factories. The proposed model enables collaborative integration for resource and demand management. Li and Sheng (2011) proposed a multi-agent model that integrates a price negotiation system based on the use of intelligent agents capable of processing information uncertainty. The model by Ogier et al. (2013 and 2015) is suitable for planning production, transport and storage activities in a supply chain at the tactical level for a limited horizon.

Short-term SC planning problems are the subject of the following work. Huang et al. (2006) offered an extension to affected operations rescheduling (AOR). Enjalbert et al. (2011) developed a distributed simulation tool for the feasibility assessment of multisite scheduling. Blackhurst et al. (2005), in their work, analyzed flexible reconfiguration tools for a supply chain in real time. Chan and Chan (2009) examined the impact of negotiation-based information sharing in a distributed contract manufacturing supply chain in a multi-period and multi-product environment modelled as a multi-agent system. Confessore et al. (2008) proposed a quick calculation method for the pricing of a delivery service in a logistics network. The goal was to develop a supportive method for effective interaction between logistics service providers and customers. Mishra et al. (2012) presented a multi-agent architecture for recycling and reverse logistics and considered aspects such as waste classification, recycling, logistics and the reuse of products.

Zhang et al. (2011b) and Blos et al. (2018) dealt generally with MASs in SCM. A review and classification of Petri net applications for SCM was provided by Zhang et al. (2011b), and Blos et al. (2018) presented a framework for modelling, analyzing, controlling and monitoring a robust and global SC and designing and integrating models with Petri networks and agent-based models for incident management.

4 Conclusion

This work provides a detailed insight into the state of the art of agent-based modelling and simulation concepts and outlines the use of MASs in the context of supply chain management. The analysis of the literature was carried out in two stages. The level 1 results demonstrated that agent-based modelling and simulation techniques are used to represent social, economic, environmental, medical, business and related systems to provide computer-assisted problem solving or cognitive enhancements.

The research performed in the IJPR in stage 2 clearly showed that the application possibilities with a focus on the area of production and logistics are very diverse.

Based on this analysis, the studies focusing on SCM were selected and classified in the supply chain planning matrix. Significant gaps in the coverage of the individual planning horizons and processes came to light. In the case of long-term planning tasks, this particularly concerns the selection of a production system, site planning, the design of a distribution network and product programme as well as strategic sales planning. In the medium term, distribution planning and tactical sales planning were ignored, as was sales planning at the operational level. The reasons for this could be the type of literature research with a focus on just one journal or MASs not being an appropriate way to deal with these planning problems. This should be investigated in subsequent work along with the mentioned gaps in connection with MASs that have not yet been investigated.

References

Allwood, J.M., Lee, J.-H.: The design of an agent for modelling supply chain network dynamics. Int. J. Prod. Res. **43**(22), 4875–4898 (2005). https://doi.org/10.1080/00207540500168295

Bechtel, C., Jayaram, J.: Supply chain management: a strategic perspective. The Int. J. Logist. Manag. **8**(1), 15–34 (1997)

Blackhurst, J., Craighead, C.W., Elkins, D., Handfield, R.B.: An empirically derived agenda of critical research issues for managing supply-chain disruptions. Int. J. Prod. Res. **43**(19), 4067–4081 (2005). https://doi.org/10.1080/00207540500151549

Blos, M.F., da Silva, R.M., Wee, H.M.: A framework for designing supply chain disruptions management considering productive systems and carrier viewpoints. Int. J. Prod. Res. **56**(15), 5045–5061 (2018)

Caridi, M., Cigolini, R., de Marco, D.: Improving supply-chain collaboration by linking intelligent agents to CPFR. Int. J. Prod. Res. **43**(20), 4191–4218 (2005). https://doi.org/10.1080/00207540500142134

Chan, H.K., Chan, F.T.S.: Effect of information sharing in supply chains with flexibility. Int. J. Prod. Res. **47**(1), 213–232 (2009). https://doi.org/10.1080/00207540600767764

Chan, H.K., Chan, F.T.S.: A review of coordination studies in the context of supply chain dynamics. Int. J. Prod. Res. **48**(10), 2793–2819 (2010). https://doi.org/10.1080/00207540902791843

Chen, J.-C., Wang, K.-J., Wang, S.-M., Yang, S.-J.: Price negotiation for capacity sharing in a two-factory environment using genetic algorithm. Int. J. Prod. Res. **46**(7), 1847–1868 (2008). https://doi.org/10.1080/00207540601008440

Chen, X.W., Nof, S.Y.: A decentralised conflict and error detection and prediction model. Int. J. Prod. Res. **48**(16), 4829–4843 (2010). https://doi.org/10.1080/00207540903067201

Confessore, G., Corini, D., Stecca, G.: A computational method for pricing of delivery service in a logistics network. Int. J. Prod. Res. **46**(5), 1231–1242 (2008). https://doi.org/10.1080/00207540701224285

Cooper, H., Valentine, J.C.: Research synthesis and meta-analysis. In: Handbook of Research on Adult Learning and Development, pp. 184–202. Routledge (2008)

Dai, H., Lin, J., Long, Q.: A fractal perspective-based methodological framework for supply chain modelling and distributed simulation with multi-agent system. Int. J. Prod. Res. **52**(22), 6819–6840 (2014). https://doi.org/10.1080/00207543.2014.919414

Davidsson, P., Henesey, L., Ramstedt, L., Törnquist, J., Wernstedt, F.: An analysis of agent-based approaches to transport logistics. Transp. Res. Part C: Emerg. Technol. **13**(4), 255–271 (2005). https://doi.org/10.1016/j.trc.2005.07.002

Enjalbert, S., Archimède, B., Charbonnaud, P.: Distributed simulation of virtual workshops for the multi-site scheduling feasibility evaluation. Int. J. Prod. Res. **49**(22), 6663–6676 (2011). https://doi.org/10.1080/00207543.2010.520911

Ferber, J., Kirn, S.: Multiagentensysteme. Eine Einführung in die Verteilte Künstliche Intelligenz. Addison-Wesley (Agententechnologie), München (2001)

Fleischmann, B., Meyr, H., Wagner, M.: Advanced planning. In: Supply chain management and advanced planning, pp. 81–106. Springer (2005)

Franklin, S., Graesser, A.: Is it an agent, or just a program? A taxonomy for autonomous agents. In: 1996 Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, pp. 21–35 (1996)

Ghirardi, M., Menga, G., Sacco, N.: An optimisation-oriented model of distributed supply-chain. Math. Comput. Simul. **79**(4), 937–946 (2008)

Huang, G.Q., Lau, J.S.K., Mak, K.L., Liang, L.: Distributed supply-chain project rescheduling: part II—distributed affected operations rescheduling algorithm. Int. J. Prod. Res. **44**(1), 1–25 (2006). https://doi.org/10.1080/00207540500151507

Huang, G.Q., Lau, J.S.K., Mak, K.L.: The impacts of sharing production information on supply chain dynamics: A review of the literature. Int. J. Prod. Res. **41**(7), 1483–1517 (2003). https://doi.org/10.1080/0020754031000069625

Jain, V., Wadhwa, S., Deshmukh, S.G.: Select supplier-related issues in modelling a dynamic supply chain: potential, challenges and direction for future research. Int. J. Prod. Res. **47**(11), 3013–3039 (2009). https://doi.org/10.1080/00207540701769958

Klügl, F.: Multiagentensystem. In: Görz, G., Schneeberger, J. (eds.) Handbuch der künstlichen Intelligenz, pp. 527–556. Walter de Gruyter (2012)

Lee, H.L., Billington, C.: Material management in decentralized supply chains. Oper. Res. **41**(5), 835–847 (1993). https://doi.org/10.1287/opre.41.5.835

Li, J., Sheng, Z.: A multi-agent model for the reasoning of uncertainty information in supply chaleeins. Int. J. Prod. Res. **49**(19), 5737–5753 (2011). https://doi.org/10.1080/00207543.2010.524257

Long, Q.: Distributed supply chain network modelling and simulation: integration of agent-based distributed simulation and improved SCOR model. Int. J. Prod. Res. **52**(23), 6899–6917 (2014). https://doi.org/10.1080/00207543.2014.910623

Lu, T.P., Chang, T.M., Yih, Y.: Production control framework for supply chain management—an application in the elevator manufacturing industry. Int. J. Prod. Res. **43**(20), 4219–4233 (2005)

Richard Martin, P., Patterson Wayne, J.: On measuring company performance within a supply chain. Int. J. Prod. Res. **47**(9), 2449–2460 (2009). https://doi.org/10.1080/00207540701725604

Mishra, N., Kumar, V., Chan, F.T.S.: A multi-agent architecture for reverse logistics in a green supply chain. Int. J. Prod. Res. **50**(9), 2396–2406 (2012). https://doi.org/10.1080/00207543.2011.581003

Nwana, H.S., Lee, L.C., Jennings, N.R.: Coordination in software agent systems. Br. Telecom Tech. J. **14**(4), 79–88 (1996)

Ogier, M., Chan, F.T.S., Chung, S.H., Boissière, J.: Decentralised capacitated planning with minimal-information sharing in a 2-echelon supply chain. Int. J. Prod. Res. **23**(16), 4927–4950 (2015). https://doi.org/10.1080/00207543.2015.1005763

Ogier, M., Cung, V.-D., Boissière, J., Chung, S.: Decentralised planning coordination with quantity discount contract in a divergent supply chain. Int. J. Prod. Res. **51**(9), 2776–2789 (2013). https://doi.org/10.1080/00207543.2012.737951

Parunak, H.V.D.: Agents in overalls: experiences and issues in the development and deployment of industrial agent-based systems. Int. J. Coop. Inf. Syst. **9**(03), 209–227 (2000)

Rohde, J., Meyr, H., Wagner, M.: Die supply chain planning matrix. Darmstadt Technical University, Department of Business Administration (2000)

Rosas, J., Camarinha-Matos, L.M.: An approach to assess collaboration readiness. Int. J. Coop. Inf. Syst. **47**(17), 4711–4735 (2009). https://doi.org/10.1080/00207540902847298

Shukla, N., Kiridena, S.: A fuzzy rough sets-based multi-agent analytics framework for dynamic supply chain configuration. Int. J. Coop. Inf. Syst. **54**(23), 6984–6996 (2016). https://doi.org/10.1080/00207543.2016.1151567

Wang, X.H., Wong, T.N., Wang, G.: Knowledge representation for multi-agent negotiations in virtual enterprises. Int. J. Coop. Inf. Syst. **49**(14), 4275–4297 (2011). https://doi.org/10.1080/00207543.2010.518996

Wooldridge, M.: An Introduction to Multiagent Systems. Wiley (2009)

Xu, L.D.: Information architecture for supply chain quality management. Int. J. Coop. Inf. Syst. **49**(1), 183–198 (2011). https://doi.org/10.1080/00207543.2010.508944

Yu, C., Wong, T.N., Li, Z.: A hybrid multi-agent negotiation protocol supporting supplier selection for multiple products with synergy effect. Int. J. Coop. Inf. Syst. **55**(1), 18–37 (2017a). https://doi.org/10.1080/00207543.2016.1189105

Yu, Y., Cao, R.Q., Schniederjans, D.: Cloud computing and its impact on service level: a multi-agent simulation model. Int. J. Coop. Inf. Syst. **55**(15), 4341–4353 (2017b). https://doi.org/10.1080/00207543.2016.1251624

Zhang, L., Wang, S., Li, F., Wang, H., Wang, L., Tan, W.: A few measures for ensuring supply chain quality. Int. J. Coop. Inf. Syst. **49**(1), 87–97 (2011a). https://doi.org/10.1080/00207543.2010.508965

Zhang, X., Lu, Q., Wu, T.: Petri-net based applications for supply chain management: an overview. Int. J. Prod. Res. **49**(13), 3939–3961 (2011b)

Dynamic Responsive Pricing as a Mitigation Strategy Against Supply Chain Disruptions: An Agent-Based Model

Niels Bugert(✉) and Rainer Lasch

TU Dresden, 01069 Dresden, Germany
niels.bugert@tu-dresden.de

Abstract. Supply Chain disruptions can result in immense financial losses for affected enterprises. Quantitative models which analyze the impact of supply chain disruptions and, in particular, the possible application of mitigation strategies can support the decision making process of practitioners to better cope with disruptions. Since existing approaches have mainly investigated the effects of backup supply and information exchange, further mitigation strategies need to be implemented. Therefore, we present an agent-based model in which the supply chain entities set their prices autonomously and dynamically based on their experienced total costs. We analyze whether dynamic responsive pricing is an appropriate strategy in the event of a disruption in case of price-sensitive customers. Our results illustrate that, in many cases, a dynamic price choice delivers better results than a fixed price choice. The value of optimal price elasticity increases the lower the price sensitivity becomes, but the speed of growth decelerates. However, if the price elasticity is too high, strong costs can occur and fixed prices become advantageous.

Keywords: Supply chain risk management · Agent-based modelling · Responsive pricing

1 Introduction

Material flow disruptions pose a serious threat to today's supply chains as risk exposure and associated risk consequences increase, particularly as a result of stronger international cooperation and an emphasis on lean management and lean logistics [1]. With nearly 1,700 incidents recorded in 2017, the automotive industry, which was the most affected industry that year according to JLT's Automotive Supply Chain Disruption Report 2018, saw a 30% increase from the 1,300 incidents reported in 2016. The most common incidents were plant fires, mergers and acquisitions, hurricanes, and labor strikes. The severity of the damage is evident from the mean recovery time, which averages 52 days for strikes, 22 days for floods, and 22 days for power outages [2]. In supply chains, the effects of disruptions can spread across the system's links, causing other entities to become impaired, thereby multiplying the damage. This effect is called the ripple effect [3].

The academic field of supply chain risk management has attracted considerable interest over the last ten years due to the severity of the damage caused by disruptions.

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 75–88, 2019.
https://doi.org/10.1007/978-3-030-29821-0_6

Intensive research has been undertaken to provide empirical research, conceptual theories on risk and risk reduction, and quantitative models for assessing risks and their impacts on the entire supply chain. The latter also integrates mitigation strategies to test their effectiveness and motivate practitioners to integrate beneficial strategies. Among the most popular mitigation strategies are backup and contingent supply as well as information sharing among partners. Even though some mitigation strategies have been studied profoundly, there are still a number of mitigation strategies that have not been considered [4].

Changing prices regarding supply chain disruptions which is also called responsive pricing is considered to be a promising reactive mitigation strategy against disruptions. [5, 22, 24]. Pricing strategies in supply chains with more than three entities have not yet been investigated in simulation models, except for Bugert and Lasch [5], who modeled responsive pricing as a reactive strategy against disruptions in a supply chain with two substitutable products and price-sensitive demand. This is rather surprising, as flexible pricing today is being facilitated by much higher data availability, information systems, more flexible pricing contracts, and e-commerce [6]. Amazon, for example, is reported to change its prices 2.5 million times a day [7].

The main objective of this paper is to present a model of a four-tier supply chain to analyze price elasticity in the presence of a significant material flow disruption and price-sensitive demand. The dynamic price changes occur once the supply chain is hit by the disruption and return to the steady state once the supply chain has totally recovered. A detailed description of the model's logic and parameters will facilitate implementation for practitioners to investigate price elasticity individually.

The rest of the paper is structured as follows. Section 2 gives a brief overview of quantitative approaches which simulate a supply chain with more than three participants and implement mitigation strategies. The outlines of the research methodology of this approach and the presentation of the corresponding research questions are presented in the third section. Section 4 presents the simulation model and discusses in detail the assumptions of the model, the model logic, and the model parameters. In Sect. 5, the results of the simulation experiment are illustrated and the presented research questions answered. Section 6 gives a brief outlook for future research work.

2 Literature Review

Several simulation techniques, such as Petri-Nets (PNs), System Dynamics (SD), Discrete Event Simulation (DES), and Agent-based Modeling (ABM), are widely used to model the dynamic impact of supply chain disruptions and the effects of mitigation strategies [5, 8–10]. In this section, we examine simulation models that concentrate on the dynamic modeling of mitigation strategies in supply chains with at least three entities. A detailed overview of supply chain disruption models can be found in Bugert and Lasch [5], while the ripple effect is further investigated in Dolgui et al. [11].

Tuncel and Alpan [12] examine the impact of mitigation strategies on a four-tier supply chain based on three disruption risks, namely quality risks, transport risks, and process risks by using a PN. Risk mitigation is modeled conceptually as a reduced probability of occurrence. Kano et al. [13] study the recovery phase after a disruption in a three-tier supply chain. With their PN, the authors investigate the influence on the supply chain's productivity assuming that a backup supplier is identified after a certain period of time. Zhang [14] tests the effectiveness of information sharing by using a PN in a three-tier supply chain with two different products if stock-outs occur. The authors differentiate between five different customer behavior patterns during stock-outs.

Wilson [8] uses an SD approach to analyze the impact of a ten-day transportation disruption at multiple locations on a five-tier supply chain. A traditional supply chain and a supply chain in which demand information is shared are compared taking into account stock levels and service levels. Sidola et al. [15] compare the impact of two transport disruptions on a regular and a visible supply chain, where all demand information is exchanged between supply chain partners. Performance metrics include inventory variability and the number of stock-outs of the retailer. Wang et al. [9] consider a two-stage supply chain with one retailer and two independent suppliers, one of which is affected by a disruption. In their SD approach, the authors compare the use of a contingent supplier, who only receives orders in the event of a disruption, with a standby supplier, who maintains a certain capacity at a regular price and charges a higher price in the presence of higher demand. Li et al. [16] explore the impact of 13 different disruption risks on the performance of a chemical supply chain with an SD model and evaluate the impact of increased transport capacity. Keilhacker and Minner [17] examine the impact of five mitigation strategies (product substitution, recycling, increased research and development, and a mix of these strategies) to address the specific problem of supply shortages of critical earth elements due to export restrictions. Using extensive empirical data, an end-to-end supply chain model is developed, consisting of a large number of mining companies, raw material processors, manufacturers, and research laboratories. Bugert and Lasch [5] model a five-tier supply chain with two products in order to examine the consequences of supply chain disruptions of a different length and model the effect of responsive pricing as a mitigation strategy.

The DES model by Schmitt and Singh [18] considers a three-tier supply chain with two suppliers, a central packaging plant, and two distribution centers with two products and predefined mitigation strategies. The model is combined with a Monte Carlo simulation to determine an aggregated distribution of the frequency and duration of disruptions per site. The authors investigate the impact of different inventory levels on the supply chain. Aqlan and Lam [10] combine Goal Programming with a DES model to find the optimal mitigation strategies, inventory levels, and production volumes under budget constraints in a supply chain with four suppliers. Mitigation strategies are modeled by an abstract reduction of risk values. In the DES model of Ivanov [3] a four-tier supply chain with realistic transport distances is presented. Different recovery and disruptive scenarios are examined to analyze the ripple effect. Ivanov [19] further studies the ripple effect as well as the post-disruption periods with a DES model. The author examines the influence of the network design and mitigation strategies, such as flexible capacity and backup supply, in a real-life case study with a multitude of performance factors.

Seck et al. [20] present an ABM approach to study the effect of different risk and recovery scenarios on the fill rate of a three-tier supply chain system with two suppliers and two sub-suppliers. The ABM model of Ledwoch et al. [21] compares a random and scale-free network topology with a single original equipment manufacturer and 102 supplier nodes with respect to different disruption frequencies and durations. The fill rate, backlog, and inventory holding costs are considered as evaluation criteria. A random number of producing entities can perform two mitigation strategies, namely contingent rerouting by transferring orders to unimpacted suppliers and increasing buffer inventory.

Flexible sourcing [9, 13, 19, 21] and information sharing [8, 14, 15] are the dominant mitigation strategies modeled in simulation models regarding supply chain disruptions. Other commonly modeled strategies are increased transport or production capacities [16, 19], buffer inventory [18, 21], abstract mitigation by lowering the probability of risks [10, 12], and modeling disruption recovery [3, 20]. Apart from these strategies, only Bugert and Lasch [5] as well as Keilhacker and Minner [17] model different strategies like static responsive pricing, recycling, etc.

In theoretical contributions within the research area of supply chain risk management, a multitude of mitigation strategies are recommended, ranging from abstract strategies such as risk acceptance, risk avoidance, risk reduction, and risk transfer to a multitude of specific approaches. Rajesh et al. [22] summarize, for example, 21 risk mitigation strategies, such as silent product rollovers, standardization, process postponement, flexible supply contracts, etc. Further publications with a summary of a multitude of mitigation strategies can be found in Aqlan and Lam [23] and Tang [24]. As dynamic, autonomous responsive pricing as a measure against supply chain disruptions has not yet been investigated, we propose a model that quantifies its usefulness in a four-tier supply chain.

3 Research Methodology

Supply chains can be regarded as complex, dynamic systems with a multitude of stochastic influencing factors. The higher the complexity of the system, the more suitable simulation becomes compared to analytical methods. We decide to conduct a simulation experiment in order to gain a deeper understanding of the system's behavior under purposefully varying input parameters [25].

Several simulation techniques with unique features and applications are available, such as System Dynamics, Petri-Nets, Agent-based Modeling, Monte-Carlo Simulation, etc. Agent-based Modeling is of particular interest for our scope of research, as it is able to represent self-organizing systems in which self-directed entities interact and influence each other [26]. In our model, the agents set their daily prices individually and interact with their direct partners via information and material flows.

We have checked the plausibility of our model using extreme condition tests. For this purpose, we set the length of the disruption to values of 0 and 50 days, increased demand to 500 and reduced it to 100 pieces per day, set the target stocks to 20,000 pieces, and then check the inventories and costs incurred.

Pre-experimental planning, the definition of research goals, the choice of factors, levels, and ranges of variables, and the selection of response variables are considered to be of importance, as they influence the choice of the experimental design [27]. Our research questions (RQs) are defined as follows:

- RQ1: How do fluctuating prices perform compared to fixed prices if a disruption occurs?
- RQ2: How should the degree of price elasticity be chosen to optimize overall supply chain profit when demand is price-sensitive?
- RQ3: How does optimal price elasticity affect each partner's profit?

The definition of the three RQs has helped us to configure the experiment and define relevant factors which pose as input variables for the simulation experiment [28]. The degree of price elasticity and the customers' price-sensitivity are designated as the relevant factors which are varied in the experiment. To answer the first RQ, we will also use fixed prices in one scenario and compare its performance to variable prices.

In simulation studies, optimization is generally achieved by developing a response surface that represents the approximate relationships between factors and responses. Since a complete calculation of the responses of all possible factors would result in an incalculable computational effort, various designs, such as full and fractional factorial design, finer grids, space-filling designs, etc., are used to explore the response surface with reasonable effort [28]. We chose a uniform space-filling design where the examined factor-response signals are evenly distributed in the factorial range. Prices are defined by the agents based on historical data about their total costs and their desired yield percentage. The moving average of the total costs is apportioned to one product sold, and the profit percentage is added. The degree of price elasticity can be adjusted with the order of the moving average, which is determined to be the first factor and referred to as the price smoothing parameter. It will vary between 30 and 150 in steps of one. If the order is smaller than 30, the system becomes unstable. A normally distributed maximum price is defined for each customer up to which the customer is willing to buy. If the sales price exceeds this maximum price, the customer refuses the purchase. The mean of this maximum price is our second factor and varies between \$30 and \$150 in steps of five.

For a variable price selection, the system requires between 400 and 1000 days until it reaches stability. The 20-day disruption of the producer therefore occurs on day 1500. To ensure that the simulation runs until equilibrium is restored, the simulation runs until day 2500.

Performance variables are the total supply chain profit and the profit of each supply chain member. To take into account the stochastic properties of the model, each experiment iteration will be replicated 25 times, and the mean of the replications' performance serves as the iteration's result. The response surface is constructed by interpolating the calculated values with Kriging, an interpolation algorithm which estimates the desired value as a weighted sum of data values of the surrounding locations. The weights are optimized using a fitted variogram function [29].

4 Model Description

4.1 Model Assumptions

The presented model consists of producing and non-producing entities in a four-tier linear supply chain. The daily number of potential customers is normally distributed.

Each customer is assumed to buy a single product if the current retail price is below their maximum price expectation and the customer's waiting time limit has not been exceeded. The customer's maximum waiting time and accepted price level are stochastically modeled by a Gaussian distribution. Unmet retailer demand is considered as lost sales and is valued with opportunity costs. Upstream entities, namely the wholesaler, the producer, and the producer's supplier, are penalized with backlog costs for every piece which they are not able to dispatch directly. The backlog costs per piece are dependent on their current sales prices.

Inventory levels of producing entities are differentiated between raw material, work in progress, and finished products. The value per piece of raw material is determined by the mean purchase price, while finished products are valued by the current sales price. The value of a semi-finished good in the production process is considered to be exactly in the middle between the purchase and the sales price. Inventory levels of non-producing supply chain members are not differentiated into different categories and valued with the mean of the purchase price and the current sales price. Holding costs per piece are incurred due to capital lock-up and warehousing costs. The inventory levels are replenished by a classical order-up-to policy. To guarantee a steady flow through the supply chain and relatively steady flexible prices in the system's balanced state, it is assumed that the stock levels are checked daily and the order points are defined in a way that every entity orders once a period if the system is in balance. The order policy is not optimal for the multi-tier supply chain. Fixed and variable order costs are tracked, and orders are immediately passed upstream.

If inventory is available upstream, the order amount is dispatched once a day and received by the downstream partner after a fixed individual transport time has passed. It is assumed that a service provider carries out the transports from entity to entity, so that the transport capacities are flexible an easily scalable. The transport costs per product are assumed to be negligible due to economies of scale.

Production processes require a uniformly distributed time span to occur and take place at the producer and his supplier. A set production capacity limits the amount of work in progress. Once goods are available in the inbound warehouse, production is initiated. Fixed and variable production costs are associated with the production.

The total resulting costs and the amount of sold goods are tracked daily and documented. Each entity calculates its own price so that a predetermined profit margin is achieved. The moving average of the total costs and goods sold together with the profit margin are used to calculate the current price, which is updated daily. The retailer is the only entity which uses a demand forecast for its price calculation to prevent the system from reaching an unstable state. The price elasticity can be adjusted by changing the length of the time window considered for the moving average. The shorter the time window, the more elastic the prices behave.

A 20-day disruption occurs at the producer's facility after the system has reached its balance. The disruptive event could be, for example, a breakdown of the IT system or the production infrastructure. During the disrupted period it is not possible for the producer to send out goods, to produce, or to order supply. Stock levels are frozen but not damaged. Goods which already have been ordered and downstream orders can be received. The profit and loss calculation still takes place each day and prices are adjusted.

4.2 Agent Organization

An agent is an autonomous, self-directed, individual entity which can function independently from other agents [26]. Each supply chain entity is modeled as a specific agent. All four agents have individual characteristics as well as common behavior.

Each agent is able to send out shipments and receive orders from the downstream partner or end customer. The current prices are sent out to each downstream entity daily. Every agent calculates its sales price, compares the received orders with the inventory, and, if possible, dispatches the total required order amount. If the inventory is too low for a complete delivery, the order is partially fulfilled and will be finished once inventory is available. Each agent tracks important data such as the inventory levels, the amount of sold goods, and the backlogged order amount. After all satisfiable orders have been dispatched, the current inventory level and the expected quantity from the upstream partner are compared with the predefined target inventory level to calculate the order amount. Orders will be sent out instantly, and the expected volume from the supplier will be increased by the order amount and decreased once the shipment has arrived. The associated fixed and variable order costs, which are dependent on the current price of the upstream partner, are tracked. After the order process has been completed, the current service level is calculated and stored. The retailer calculates the service level based on the delivered order amount of this day and the total desired order amount of this day. All other entities consider the ratio of the instantly fulfilled order amount to the total desired order amount of this work day. The last daily process step of each entity is to calculate the profit. All relevant costs such as order costs, production costs, holding costs, and backlog costs are subtracted from the current revenue, and the cost and profit data is stored.

The producer and their supplier are the only entities with a production process. The producer's supplier is placed at the upstream end of the modeled supply chain. Therefore, his supplier issues a fixed price for raw material, and the ordered material appears in the inbound warehouse of the producer's supplier after a predefined delivery time. The production process of these two entities is initiated once inbound material is available and starts after the order process has finished.

4.3 Model Parameters

The number of daily customers which desire to buy a single product is modeled by a normal distribution with a mean of 250 pieces per day and standard deviation of 10 pieces per day. Customers are willing to wait a maximum of five days for their product if the retailer's inventory is empty. Holding costs, which include the physical warehousing costs and capital lockup costs, are set at 18% of a product's value per year, which results in 0.05% per day if a 360-day commercial year is considered.

The variable ordering costs correspond with the current forwarded price of the upstream entity. The price of the raw material at the upstream end of the supply chain is considered to be $10 per piece. The maximum production capacity of the supplier and the producer is set at 2000 pieces. The yield per piece was set at 0.0% to simplify the quantification of the disruption costs. Remaining parameters can be found in Table 1.

Table 1. Model parameters.

	Supplier	Producer	Wholesaler	Retailer
Minimum production time [days]	4	4	–	–
Maximum production time [days]	6	6	–	–
Production setup costs [$]	50	50	–	–
Variable production cost [$/piece]	2.5	5	–	–
Fixed order cost [$]	90	180	120	90
Target inventory [pieces]	6000	3500	2000	1500
Delivery time [days]	6	4	2	–

5 Discussion of Results

5.1 Disruption with Fixed Prices

We first consider a supply chain disruption with fixed prices as the first part to answer the first RQ, so that the desired yield percentage of 0.0% is achieved by each member. This yield percentage allows us to determine the disruption costs with relative ease because the profit stays comparatively constant after the system has stabilized again. The mean value of the profit over 100 days is used as a performance measure. When the system is in balance, prices fluctuate only slightly, so the average prices over 100 days are selected. For the retailer this results in a retail price of $20.00, for the wholesaler $19.62, for the producer $19.11, and for the supplier $13.13. It is assumed that all customers accept the sales price, so that the service level in a balanced system state is at 100%.

Figure 1 presents the inventory fluctuations caused by a 20-day disruption of the producer after 1500 days. It is worth noting that due to the stochastic properties of the system, 25 replications are carried out for a fixed price. Therefore, Fig. 1 shows the results of only one replication. Other replications lead to slightly different values, but have the same basic properties.

As soon as the disruption occurs, the producer can no longer process orders and therefore cannot dispatch purchasing orders or shipments. Purchasing orders that have already been placed at the supplier still arrive within the next six days and are received. The inventory level first increases and remains constant for the duration of the disruption, before it is rapidly emptied by the accumulated orders that could not be executed. After the inventory drops to zero, all produced goods are shipped out instantly, leaving the inventory level close to zero until day 1570, followed by a gradual recovery to normal conditions. The ripple effect can easily be identified by the fact that this disruption first affects the wholesaler and subsequently the retailer around five days later. The warehouses of both members are emptied and then recover to normal levels. Since no orders are passed on from the producer to the upstream entity, the supplier's inventory quickly rises to 4000 before rapidly declining and finally recovering. The supplier is therefore the only member not to experience a stock-out.

Figure 2 shows that as soon as the effects of the disruption are felt completely, the service level of the producer drops considerably. Since the service level is calculated

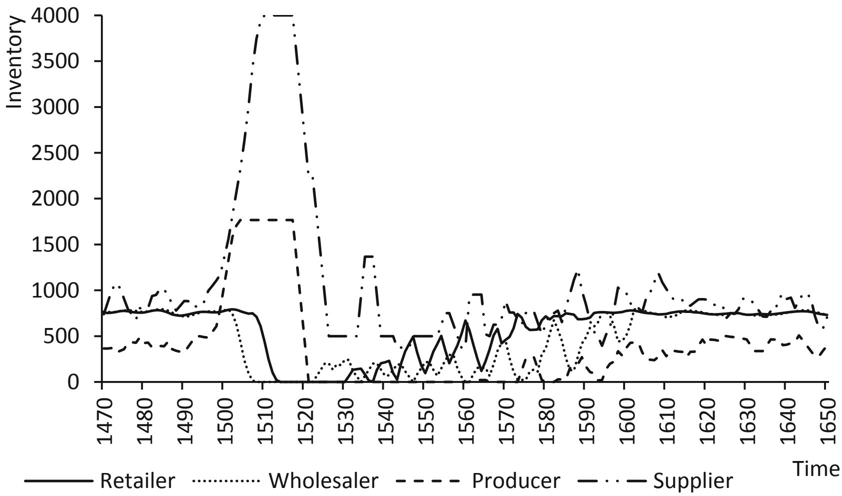


Fig. 1. Inventory fluctuations caused by the 20-day disruption.

based on data from the last 10 sales days, it reaches 0 after 10 days. It then remains at this level until the direct effect of the disruption ceases on day 1520. The entire inventory is dispatched instantly and all items produced are sent out immediately. The service level remains at around 30% to 50% for another 40 days until it slowly recovers. The producer reaches an optimal service level around day 1620. With a time delay of four days, the wholesaler's service level also drops abruptly and slowly recovers from day 1525 onwards. Approximately 100 days after the disruption has started, the service level of 100% is reached again. The retailer can stay at a service level of 100% for an extended time, while the supplier's service level stays unaffected at 100%, as there is enough inventory available. The impact is felt in the supply chain for more than 120 days.

The resulting disruption costs can be seen in Table 2. The producer experiences the worst effects with an average of $1,008,206. He has to struggle especially with the backlog costs for a prolonged period of time. The wholesaler initially has high storage costs due to the lack of sales and then has to accept additional backlog costs. The supplier has to bear heavy storage costs. The retailer is affected for a relatively short time, but is affected by the disruption due to lost sales. Overall, the supply chain is severely affected with average losses of $1.68 million. The range of losses can be explained by the model's high stochastic influences.

5.2 Disruption with Variable Prices

This section examines the extent to which variable prices are able to mitigate the financial consequences of disruptions in order to address the second and third RQ. First, we want to illustrate how the agents set prices in the event of a disruption. A price elasticity of 60 and an average maximum customer price of $85 were selected. Figure 3 shows the price behavior during the 20-day malfunction of the producer starting on day

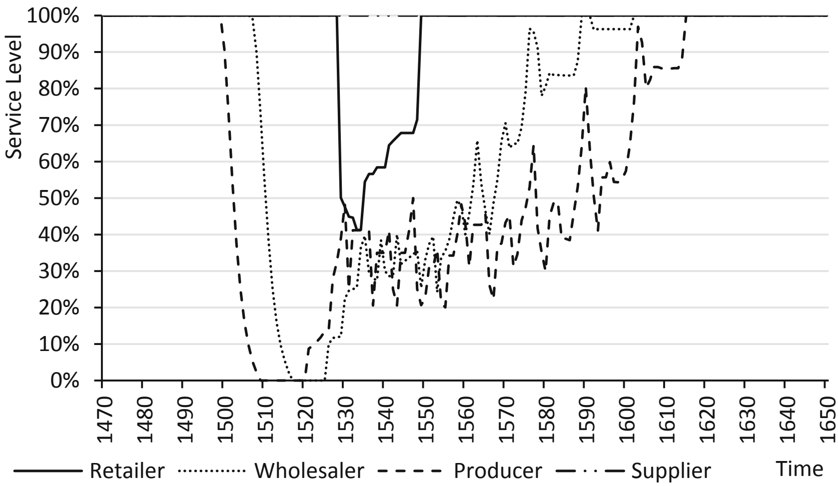


Fig. 2. Changes in service levels caused by the disruption.

Table 2. Disruption costs in case of fixed prices.

Disruption costs	Total	Retailer	Wholesaler	Producer	Supplier
Average [$]	1,682,213	48,500	442,480	1,008,206	182,972
First quartile [$]	1,622,895	47,684	411,042	968,107	170,699
Median [$]	1,652,762	48,353	442,829	990,215	183,502
Third quartile [$]	1,746,650	49,196	463,091	1,044,973	193,225

1500. Since the producer is the first to be affected by the disruption, the price rises sharply and reaches its peak at $95 on day 1610. After that, the price needs until about day 1770 to stabilize again. After the prices of the producer rise, the wholesaler is hit by the disruption and reacts with price increases amounting to a maximum of $100. Through the ripple effect, the consequences of the disruption reach the retailer after a longer time lag. The retailer price climbs to around $85 and then returns to its initial level. Although the delivery problems and stock levels have been resolved by day 1620, prices take a considerably longer time to rebalance.

The autonomous price adjustment has an impact on the supply chain, whose total losses decrease on average over 25 replications to $332,017. In this instance, the price adjustment is therefore preferable to a fixed price selection, as the supply chain could reduce costs by $1,350,196. Especially the producer and their supplier are able to benefit from the price adjustment and even profit from it. The producer is the biggest beneficiary and increases their profit by an average of $131,494, while the supplier experiences an increase of $21,933. The wholesaler has to cope with a loss of $73,608, which is still significantly less than the losses of a fixed price system. The retailer is hit by the high prices upstream and cannot compensate for the losses. They post a loss of $411,784 on average.

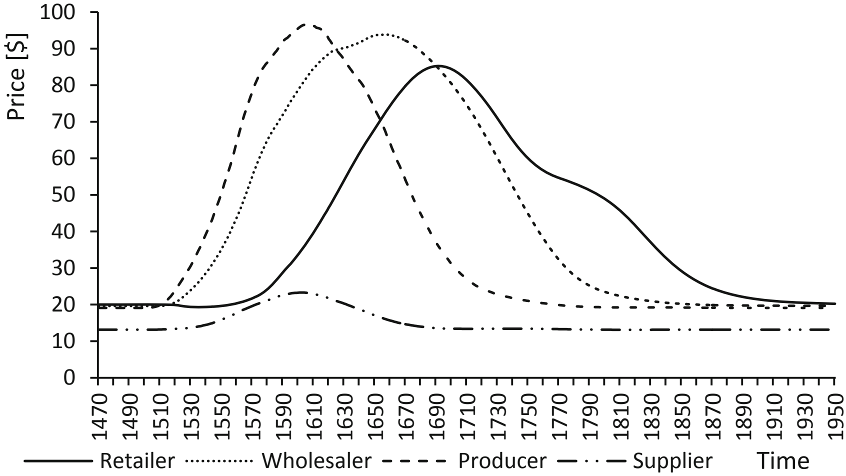


Fig. 3. Price fluctuations caused by the disruption.

The total profits of the supply chain from the parameter variation experiment are illustrated in Fig. 4, varying the price smoothing parameter as a measure for the price elasticity and the price expectancy. Referring to the second RQ, it can be observed that in a number of cases dynamic pricing leads to better results than a fixed price. It is even possible that the entire supply chain generates a profit from the disruption. At a high price elasticity with a value moving towards 30, the losses increase significantly. If the price acceptance of the customers increases, this effect can be compensated for, but if the price elasticity is too high, it cannot lead to a benefit compared to fixed prices. At a price elasticity between 60 and 90, the profit first decreases with a higher price acceptance before it sharply increases to its peak. The reason why a low price acceptance can even have positive effects stems from the fact that the retailer sells very little and the lost sales tend to be lower than the high ordering costs caused by the wholesaler's high price. Even with overly high price expectations, the total profit decreases again, but remains at a high level. With a low price elasticity and thus values approaching 150, an optimum can be found if the price acceptance reaches values of 40. It is evident that the optimum elasticity value rises with an increasing price acceptance, but excessive elasticity leads to immense costs.

Table 3 shows the optimal values for the price elasticity and the resulting profit of each entity and the total supply chain in relation to different price expectation levels to help answer the third RQ. For price expectations below $60, the optimum value lies at the edge of the range considered, with a price elasticity of 150. It can be noted that the optimum price elasticity decreases and the overall profit in the supply chain increases as customers' price expectations increase. The higher the willingness to purchase the products, the more revenue can be generated, and high prices are not penalized by the customer. However, it is interesting to see that the higher the price expectation is, the lower the absolute changes in optimal price elasticity become. Excessive elasticity damages profits in the supply chain as a strong abrupt price increase has a significant

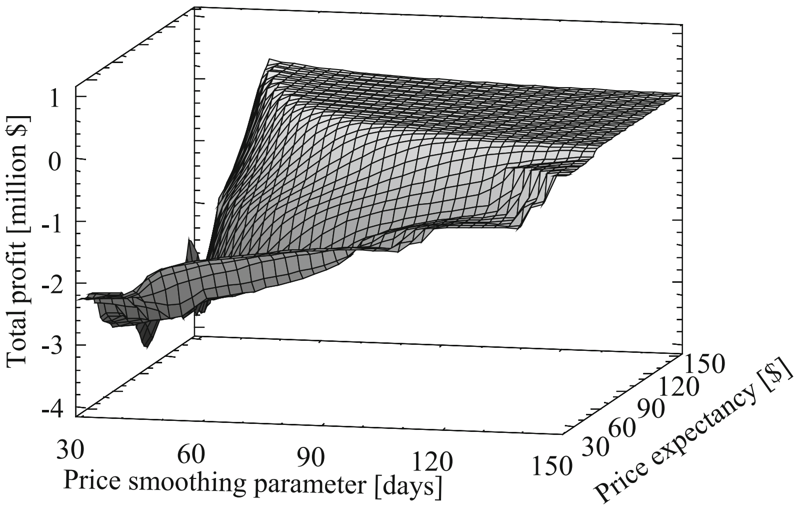


Fig. 4. Total profit with respect to price elasticity and price expectancy.

impact on other partners, causing the retailer to suddenly lose sales and the system to lose balance, resulting in higher costs.

Table 3. Resulting profits for optimal price elasticity.

		Profit [$]				
Price expectancy	Opt. smoothing parameter	Total	Retailer	Wholesaler	Producer	Supplier
60	123	97,515	−57,381	42,547	99,758	13,949
70	101	142,271	−53,799	53,224	127,455	17,830
80	83	182,475	−60,198	64,166	155,315	21,210
90	73	218,161	−61,039	75,143	178,454	24,223
100	66	251,228	−60,538	85,616	199,886	27,884
110	61	283,316	−59,553	94,622	219,647	29,720
120	57	311,192	−58,520	102,153	235,382	30,296
130	53	339,265	−60,694	111,938	256,794	32,647
140	50	364,242	−64,728	118,142	276,181	33,067
150	48	397,603	−60,096	128,041	295,468	34,511

When comparing the disruption costs of a fixed price selection with the results of the variable price selection, it can be seen that the partners most affected by the disruption profit the most from a variable price selection. The higher the disruption costs have been with fixed prices, the higher the profits are with an optimal price variance. The retailer's profit remains at a constant negative level of around $60,000.

Revenues from higher sales are lost to upstream partners due to high ordering costs. The other three entities benefit from falling price sensitivity among customers, especially the producer.

6 Conclusion

We have presented an agent-based model of a four-tier supply chain to examine whether an autonomous price adjustment based on experienced costs makes sense for each supply chain member and the total supply chain against an occurring disruption. A moving average of the total costs with different lengths serves each participant to determine the price. In this model it could be illustrated that a price adjustment is in many cases preferable to a fixed price choice. The optimal price variability increases with increasing price acceptance by the customers, but the increase becomes increasingly slower. Excessive price variability results in high costs, as the system becomes more imbalanced due to rapid price increases and the resulting loss of customer orders. In the case of an optimal variability, the entire system can even profit from a disruption. The retailer is worse off due to the price change, while all other entities profit from it. The producer who experiences the disruption directly can benefit most from the price change.

Future research could consist of finding out how the system behaves in a supply network and when prices are selected cooperatively using cost information from the entire supply chain. It could also be investigated to what extent different customer profiles with different price sensitivities influence the results. Real data from supply chains could be useful for further investigation. We want to motivate researchers and practitioners to adapt the model to an individual value chain and incorporate further mitigation strategies into quantitative models.

References

1. Stecke, K., Kumar, S.: Sources of supply chain disruptions, factors that breed vulnerability, and mitigation strategies. J. Mark. Channel **16**(3), 193–226 (2009)
2. JLT. https://www.jltspecialty.com/-/media/files/sites/specialty/insights-automotive/jlt-automotive-supply-chain-disruption-report. Accessed 10 Feb 2019
3. Ivanov, D.: Simulation-based ripple effect modelling in the supply chain. Int. J. Prod. Res. **56**(1–2), 414–430 (2017)
4. Bugert, N., Lasch, R.: Supply chain disruption models: a critical review. Log. Res. **11**(5), 1–35 (2018)
5. Bugert, N., Lasch, R.: Effectiveness of responsive pricing in the face of supply chain disruptions. Comput. Ind. Eng. **124**, 304–315 (2018)
6. Agatz, N., Fleischmann, M., van Nunen, J.: E-fulfillment and multi-channel distribution – a review. Eur. J. Oper. Res. **178**(2), 339–356 (2008)
7. Mehta, N., Detroja, P., Agashe, A.: Amazon changes prices on its products about every 10 minutes – here's how and why they do it, https://www.businessinsider.de/amazon-price-changes-2018-8. Accessed 10 Feb 2019

8. Wilson, M.: The impact of transportation disruptions on supply chain performance. Transp. Res. E-Log **43**(4), 295–320 (2007)
9. Wang, Y., Huang, M., Chen, J.: System dynamics modeling of backup purchasing strategies under supply disruption risks. Int. J. Serv Sci. Technol. **7**, 287–296 (2014)
10. Aqlan, F., Lam, S.: Supply chain optimization under risk and uncertainty: a case study for high-end server manufacturing. Comput. Ind. Eng. **93**, 78–87 (2016)
11. Dolgui, A., Ivanov, D., Sokolov, B.: Ripple effect in the supply chain: an analysis and recent literature. Int. J. Prod. Res. **56**, 414–430 (2018)
12. Tuncel, G., Alpan, G.: Risk assessment and management for supply chain networks: a case study. Comput. Ind. **61**(3), 250–259 (2010)
13. Kano, C., Blos, M., Junqueira, F., dos Santos Filho, D., Miyagi, P.: Modeling and analysis of global supply chain disruption based on petri net. In: 22 International Congress of Mechanical Engineering (COBEM 2013), pp. 382–390. ABCM, Ribeirao Preto, SP, Brazil (2013)
14. Zhang, X.: A petri net based simulation to study the impact of customer response to stock-out on supply chain performance. Cogent Eng. **3**, 1–16 (2016)
15. Sidola, A., Kumar, P., Kumar, D.: System dynamic methodological approach for design and analysis of risk in supply chain. In: First International Technology Management Conference (ITMC 2011), pp. 495–500. IEEE, San Jose, CA, USA (2011)
16. Li, C., Ren, J., Wang, H.: A system dynamics simulation model of chemical supply chain transportation risk management systems. Comput. Chem. Eng. **89**, 71–83 (2016)
17. Keilhacker, M., Minner, S.: Supply chain risk management for critical commodities: a system dynamics model for the case of the rare earth elements. Resour. Conserv. Recycl. **125**, 349–362 (2017)
18. Schmitt, A., Sing, M.: A quantitative analysis of disruption risk in a multi-echelon supply chain. Int. J. Prod. Econ. **139**(1), 22–32 (2012)
19. Ivanov, D.: Disruption tails and revival policies: a simulation analysis of supply chain design and production-ordering systems in the recovery and post-disruption periods. Comput. Ind. Eng. **127**, 558–570 (2018)
20. Seck, M., Rabadi, G., Koestler, C.: A simulation-based approach to risk assessment and mitigation in supply chain networks. Procedia Comput. Sci. **61**, 98–104 (2015)
21. Ledwoch, A., Yasarcan, H., Brintrup, A.: The moderating impact of supply network topology on the effectiveness of risk management. Int. J. Prod. Econ. **197**, 13–26 (2018)
22. Rajesh, R., Ravi, V., Venkata Rao, R.: Selection of risk mitigation strategy in electronic supply chains using grey theory and digraph-matrix approaches. Int. J. Prod. Res. **53**, 238–257 (2015)
23. Aqlan, F., Lam, S.: Supply chain risk modelling and mitigation. Int. J. Prod. Res. **53**(18), 5640–5656 (2015)
24. Tang, C.: Robust strategies for mitigating supply chain disruptions. Int. J. Logist. Res. Appl. **9**(1), 33–45 (2006)
25. Montgomery, D.: Design and analysis of experiments. Wiley, New York (2012)
26. Macal, C., North, M.: Introductory tutorial: Agent-based modeling and simulation. In: Proceedings of the 2014 Winter Simulation Conference, pp. 6–20. IEEE, Savannah, GA, USA (2014)
27. Kleijnen, J.: An overview of the design and analysis of simulation experiments for sensitivity analysis. Eur. J. Oper. Res. **164**(2), 287–300 (2005)
28. Sanchez, S.: Work smarter, not harder: guidelines for designing simulation experiments. In: WSC2006, pp. 47–57. IEEE, Monterey, CA, USA (2006)
29. Oliver, M., Webster, R.: A tutorial guide to geostatistics: computing and modelling variograms and kriging. Catena **113**, 56–69 (2014)

Applicability of Blockchain Technology in Scheduling Resources Within Distributed Manufacturing

Jacob Lohmer(✉)

Chair of Business Management, esp. Logistics, Technische Universität Dresden, Dresden, Germany
jacob.lohmer@tu-dresden.de

Abstract. Collaborative production networks are becoming a common setting in modern times. Issues of how and where data from IoT devices is stored and how communication between entities is conducted are not fully resolved. Decentralized approaches provide opportunities compared with centralized approaches in this context, especially in terms of trust and data security. One technology that has emerged recently is the blockchain technology. Although the potential benefits are well known, literature on blockchain in supply chain management and manufacturing is scarce. This paper presents a concept for the use of blockchain technology in distributed manufacturing. By utilizing smart contracts, manufacturing processes are executed on shared resources in distributed collaborative production networks. Benefits and risks of the proposed methodology are identified. The paper is completed by a case study with several companies operating in distributed manufacturing.

Keywords: Blockchain technology · Distributed manufacturing · Distributed scheduling

1 Introduction

Digitalization and Industry 4.0 as key trends in modern manufacturing and supply chain management (SCM) enable new technologies and lead to a comprehensive reengineering of manufacturing strategies and configurations (Kagermann et al. 2013; Rossit et al. 2018). These new configurations are also a reaction to market uncertainty and high demand volatility in recent years. Manufacturing enterprises are progressively forming collaborative production networks that enable them to share resources within the network (Lou et al. 2010; Renna and Argoneto 2011; Tao et al. 2017). This dynamic configuration in distributed manufacturing (DM) allows a flexible reconfiguration at any time according to market and production needs. However, communication and negotiation processes in production networks remain a complicated task and often contradict the need to respond quickly to market demands. Decisions where and how to store production data generated by new Internet of Things (IoT) devices as well as how to handle secure and reliable communication between IoT devices is a continuous topic of research (Makhdoom et al. 2019). A service that could significantly

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 89–103, 2019.
https://doi.org/10.1007/978-3-030-29821-0_7

increase trustworthiness in IoT data and complex supply chain configurations like collaborative manufacturing is provided by the blockchain technology. The blockchain as a protocol allows an immutable and decentralized storage of information and data (Nakamoto 2008). The open and decentralized architecture of the blockchain offers an alternative to existing centralized and decentralized systems that is worth investigating in the manufacturing environment (Li et al. 2018). Companies connected in a peer-to-peer (P2P) network could automatically execute manufacturing processes through smart contracts on a blockchain, without the need for human interaction. Although the potential benefits of using the blockchain in business are well known, literature on blockchain in SCM and DM is scarce. The goal of this paper is to present a methodology for a blockchain system that allows collaborative manufacturing in production networks. As the identified problem is yet unsolved, we employ the method of design science to create a new IT artifact and evaluate it. Design science as a method from information research ensures relevance and rigor of the results (Hevner et al. 2004). Specifically, we intend to answer the following research questions:

RQ1: Which issues persist in collaborative production networks and how could blockchain technology contribute to solving these issues?
RQ2: How could a methodology for a blockchain system be organized in a collaborative production network? What are the potential benefits and risks?

The remainder of this paper is organized as follows. First, we provide an overview of characteristics of collaborative production networks and indicate issues involved. Next, we examine the blockchain technology and demonstrate the potential the blockchain offers to fill the issues in collaborative production. Then, we propose a methodology that uses smart contracts to schedule machines as IoT devices. We discuss capabilities, benefits and risks of our methodology and show the applicability through a case study. The paper is closed with limitations, further research opportunities and a conclusion.

2 Theoretical Background

2.1 Distributed Manufacturing in the Industry 4.0 Era

Industry 4.0 as a collective term stands for a paradigm change in manufacturing aiming at a more flexible design of production processes. Autonomous entities are able to interact in decentralized architectures based on real-time information (Kagermann et al. 2013; Tao et al. 2017). Cyber-physical systems (CPS) as the main elements in new adaptive and evolving structures combine hard- and software characteristics, have a unique identity and are able to perceive and communicate with their surroundings in the smart manufacturing concept (Bartsch et al. 2018). In the wider concept of cyber-physical production systems, the shop floor is directly linked to higher decision support systems. Machines and parts hardly need human control to interact (Rossit et al. 2018). The smart systems have a more dynamic structure than current supply chains. This dynamic is also reflected in the recent tendency of companies to consider virtual manufacturing networks that combine advanced computing systems and coordination strategies. These networks are often built in conjunction with the broader concept of DM, which stands for geographically dispersed manufacturing in flexible networks. These trends are a reaction to the concentration of core competences and tough global competition (Tuma 1998; Rodríguez Monroy and Vilana Arto 2010)

Two main types of virtual collaborative networks can be distinguished: Temporary project-based cooperation and long-term collaborative networks with mainly independent companies that can join and exit the network autonomously (Tuma 1998; Srai et al. 2016). The networks are designed to support the rapid and flexible sharing and accumulation of resources as well as expertise among network members. The virtual component of the system is related to the intensity of collaboration with companies outside the organization itself (Shi et al. 2005; Tao et al. 2017).

A concept that is increasingly considered in research with regard to production structures with virtual collaboration and cooperation is cloud manufacturing (Ren et al. 2017). As a model for smart networked manufacturing, it enables customers to obtain on-demand services through a network with a shared pool of virtualized resources. Cloud manufacturing leads to a shift from production-oriented to service-oriented manufacturing, which can also cover only parts of the whole production process (Yang et al. 2014). Communication about tasks to be performed and scheduling (as the allocation of jobs on machines) are complex and time-consuming tasks that usually require a lot of coordinating activities in distributed manufacturing (Lou et al. 2010; Pinedo 2012). The emergence of IoT devices presents new opportunities here, as the coordination can be delegated to the devices themselves. Through this development, the implementation of concepts of Industry 4.0 may lead to a more dynamic multilateral cooperation in SCM and DM.

2.2 Issues in Collaborative Distributed Manufacturing Networks

The operation of collaborative DM networks is a scenario that is found primarily in capital-intensive industries such as semiconductor or additive manufacturing (Renna and Argoneto 2011). High demand volatility leads to idle times on valuable machines that are therefore shared in the network. When forming a collaborative manufacturing network, several decisions have to be made, in particular how and where to store production data generated from shared resources and how to handle secure and reliable communication between entities. In general, there are two options for data exchange and communication: centralized or decentralized IT systems. In a centralized architecture, one partner in the network aggregates and stores all production data and enables other partners to access the data. Solutions with central data processing, whether through cloud computing or standard IT systems, hold several risks and vulnerabilities: Data security regarding a centralized point of vulnerability and trust regarding data corruption are just two of many aspects (Makhdoom et al. 2019; Casino et al. 2019). It is time-consuming and costly to connect each partner to a central platform, which especially affects small and medium-sized enterprises. The availability of the system decisively depends on the resources of the central authority, which also determines the agility of the system. In addition, centralized networks can hardly be operated without inequalities and injustices (Korpela et al. 2017; Reyna et al. 2018). Relying on third parties (e.g. cloud service providers in cloud manufacturing) or mediators also reduces productivity and scalability (Yang et al. 2014; Ren et al. 2017). IoT makes these shortcomings even more visible since the machines as cyberpyhsical systems are basically capable of sensing and interacting with their environment freely. Routing this communication through a mediator in a central IT system or cloud manufacturing omits the fundamental advantage of IoT, the free communication with the environment in real time (Reyna et al. 2018).

Decentralized solutions are often used to conquer these drawbacks. Each actor makes his own decisions; common goals are achieved by a coordination mechanism. Most approaches in the literature apply game theory or negotiation mechanisms, multi-agent architectures are particularly popular (Renna and Argoneto 2011). A main disadvantage of multi-agent systems is that the entire communication between agents takes place in a sublevel and cannot be transparently tracked. Here again trust is the keyword, as existing decentralized systems often lack transparency (Li et al. 2018). The significance of trust and security is underlined by the recurrent reporting by companies in surveys that the lack of data exchange and high-quality information in networks is widespread and considered as a major issue (KPMG 2016). Only a quarter of companies assess the level of trust among ecosystem partners as 'high' according to another survey by Cognizant (Das 2017). To summarize the observations, the challenges in collaborative production networks and DM can be grouped in five main categories: Trust, integrity, availability, agility and security. Existing decentralized networks do not provide a flawless solution for these issues. At least security and agility as key characteristics in the concept of DM are best modeled through a network of distributed nodes (Kohtala 2015). This network of nodes can be managed as a P2P network using the blockchain.

3 Blockchain Technology

3.1 Fundamentals of Blockchain Technology

The blockchain is essentially a distributed ledger that is stored decentrally as a single record of data in blocks on network nodes. It is better known to the public as its utilization to perform the transactions of cryptocurrencies like Bitcoin (Nakamoto 2008). The underlying secure communication among nodes in the network uses a P2P network. The nodes are geographically dispersed and no single node is privileged. The communication is handled through transactions and elements of cryptography. An immutable and complete ledger is established that is replicated between all participants in the network. Each node synchronizes data and propagates transactions. A transaction in the ledger is issued for every communication event and consists of at least two addresses (sender and receiver), the message itself and signatures. Blocks record accruing lists of transactions that contain a hash pointer to reference the previous block and an immutable timestamp. In this way, a chronological 'chain' of blocks is established. For the first time, the blockchain provides entities with the opportunity to communicate and exchange values and information with other entities they do not need to trust without a mediating third party. The use of strong cryptographic algorithms ensures a secure exchange of information (Abeyratne and Monfared 2016; Bahga and Madisetti 2016).

Instead of a central authority or a mediator, a consensus mechanism is carried out in the network nodes to determine the current state of the ledger and reach an agreement on which transactions should be added to the ledger. Proof-of-Work (PoW), Proof-of-Stake (PoS) and Proof-of-Authority (PoA) are commonly employed consensus mechanisms (Zheng et al. 2017; Makhdoom et al. 2019). In the most prominent mechanism PoW, so-called miners try to solve a cryptographic algorithm to be able to add a new valid block to the chain and receive an award as digital currency credit. This incentive ensures constant validation and consistency of data throughout the network and a steady network performance. In the context of distributed manufacturing, PoA should be emphasized. PoA uses the computing costs in the network to regulate itself and facilitates fair participation among all partners (Casino et al. 2019). PoA uses a set of nodes that are specifically authorized (‘authorities’ or ‘validators’) to create new blocks rather than miners. Each new block and the resulting new blockchain must be signed by a majority of authority nodes. As the reputation of each authority is attached to its identity, there is an incentive for authorities to keep the transaction process running and uncompromised. PoA in its core is more secure, less costly and more performant than PoW as computational heavy mining is unnecessary and transaction acceptance latency lower. Blocks are also created at regular intervals, which increases predictability and stability (Parity 2019).

The design of the blockchain itself determines the access rights of the participants: Public, private and consortium blockchains are currently in use. Public blockchains perform best in terms of decentralization. Private blockchains are more efficient in terms of transaction costs and transaction speeds but the network managing entity is in a pervasive position and could restrict access for other participants or cause inequalities at will. This setting is commonly not favorable when compared with a single IT system regarding cost, availability or speed. In a consortium blockchain, the consensus process is managed by a preselected set of nodes. The main advantage of a consortium blockchain is the combination of features from both private and public chains like decentralization, stable transaction costs and fast transaction speeds (evan GmbH 2019).

A rapidly evolving area for blockchain utilization besides the use of cryptocurrencies for financial transactions are smart contracts. A smart contract represents a computerized transaction protocol that automatically executes a contract. It therefore facilitates the embedment of contractual clauses into scripting languages to perform blockchain functions (Szabo 1997; Hyperledger 2018). Smart contracts are trusted applications that are set up with a unique address and can be triggered by a transaction that indicates this address and satisfies the specified criteria of the smart contract (see Fig. 1). The smart contract then independently executes its functions in the network and changes the state variables depending on the inner logic. Scripting languages like Python or Solidity can be used to set up contracts that are then complied into byte code and uploaded to the blockchain (Bahga and Madisetti 2016). The most important feature of smart contracts is their inherent immutability. Smart contracts can specify clear contractual obligations which the entity that triggers the smart contract must comply with during execution (Kim and Laskowski 2017). Smart contracts are essential for the methodology presented in Sect. 4.

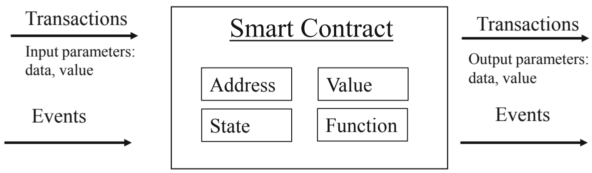


Fig. 1. Smart Contract – Structure and content (based on Bahga and Madisetti 2016)

3.2 Research on Blockchain Utilization in Supply Chain Management and Manufacturing

Research into blockchain technology and application development has started in industry itself and has experienced an upswing since 2016. Three main application areas in SCM have been distinguished so far: product tracking, product tracing and supply chain finance (Petersen et al. 2018). Real use cases in supply chain and manufacturing are rarely published in detail. IBM is noteworthy as they developed the blockchain platform 'TradeLens' to establish transparent and secure managing of documents related to maritime shipment and customs clearance (Scott 2018). The focus is on promoting industry adoption and digitalization of the global supply chain. Amazon, IBM, Oracle, Microsoft and SAP all offer their cloud services to enterprises as blockchain-as-a-service to develop their own blockchain infrastructure (Rupareliya 2018).

In contrast to the interest from industry, scientific literature on blockchain technology is still scarce. The main contributions originate from computer sciences and business informatics, mostly surveys focusing on security or privacy issues (Wang et al. 2018; Zheng et al. 2018; Feng et al. 2019). Blockchain use in SCM and manufacturing has not been addressed by research to the same extent: Petersen et al. (2018) analyzed current blockchain applications and present findings from a survey in which regulatory uncertainty and the need to join forces with different parties were seen as the toughest barriers to adoption. Bartsch et al. (2018) developed a secure additive manufacturing platform to enable networked 3D printers to automatically license products on a blockchain while ensuring traceability and data protection throughout the entire product lifecycle. Angrish et al. (2018) developed a blockchain testbed platform to share logs of machine events (process capabilities, downtimes etc.) via smart contracts in a permissioned network. Li et al. (2018) proposed an architecture to combine cloud manufacturing and the blockchain ('BCmfg'). Blossey et al. (2019) identified key features where the blockchain can improve SCM: visibility, integrity, orchestration and virtualization. Virtualization or tokenization means that claims on physical assets (or capacity slots) can be issued as 'tokens' and allow for an enhanced capacity utilization of manufacturing assets in the supply chain.

Qualitative research provides insight from industry: Korpela et al. (2017) identified system security, multi-stakeholder contracting and transactions as the most relevant areas in a study on digital supply chain integration with blockchain technology in Finland. Recent surveys by Capgemini and Staufen AG indicate that manufacturing is one of the most promising industries for blockchain adoption (Pai et al. 2018;

Staufen AG 2018). However, most professionals are still unsure about the real potential of the technology or consider it as too complex to understand and apply (Staufen AG 2018). Hence, there is still a need for clarification in this area of Industry 4.0, which adds impetus to this contribution. As several authors discovered the lack of scientific literature on the subject of blockchain technology in SCM and manufacturing, we contribute to closing this identified research gap with this paper.

3.3 Applying the Blockchain to Collaborative Production Networks

To illustrate our suggested link between collaborative production and the blockchain, see Fig. 2 (based on Pai et al. 2018). The blockchain is a promising alternative to existing centralized and decentralized approaches.

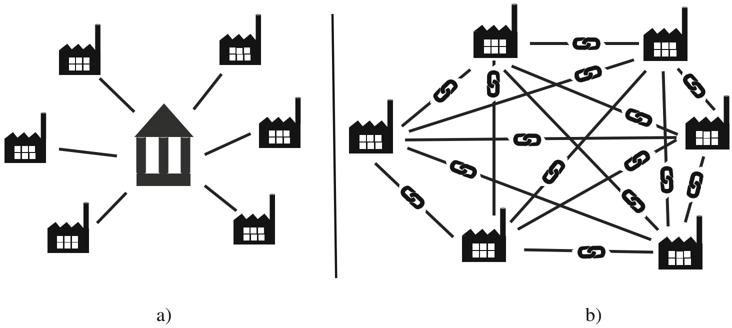


Fig. 2. Transactions in the network: centralized system (a), blockchain system (b)

Instead of connecting each network partner to a central authority (Fig. 2a), the blockchain is used to construct a P2P network among network partners (Fig. 2b). This enables a ‘radically decentralized, collaborative and nonproprietary’ environment in DM and multilateral cooperation between network partners (Benkler 2006). The use of the blockchain is accompanied by the elimination of intermediaries. This may lead to increased complexity in a collaborative network; smart contracts can be used to counter this complexity. The key technological advances of the blockchain compared to common information exchange technologies can be linked to the issues in collaborative DM as follows:

- *Security and resilience*: As a decentralized solution, the blockchain is not exposed to a single point of failure, but is maintained by all nodes in the P2P network. The data record is immutably stored. Cryptographic algorithms ensure the security of all transactions in the network. If a node is compromised and fails, the remaining nodes can compensate. The affected node can be reconfigured to quickly restore the current network state. The network can be reconfigured flexibly, entities can be added and others can leave the network easily. In contrast to central IT systems that tend to become more insecure with size, the blockchain benefits from a larger number of entities involved.

- *Transparency, integrity and immutability*: The blockchain enables a truly decentralized P2P network without the need for trust or third party mediators. No central authority can manipulate assets or transactions. Consensus is achieved through all nodes and auditing is easily possible. A validator can track the transactions and the history can be used as evidence for any legal disputes. The blockchain can also mitigate most of the liability risk between industrial partners, since all data concerning transactions is stored on the blockchain in a retrievable way with timestamp and user assignment.
- *Autonomy in decentralization*: DM profits from IoT devices that can autonomously communicate with each other using the blockchain. Smart contracts enable an automatic execution of actions inherited in code without any human interaction or tampering. Adapting a blockchain solution is also considered to influence costs; savings of about 5% are considered likely (Das 2017). Financial flow management can easily be handled incorporating a cryptocurrency.

By leveraging these features, blockchain technology can address the most challenging issues in collaborative production networks. One of many conceivable possibilities the technology offers for production environments is presented hereafter.

4 A Blockchain System for Resource Scheduling in Collaborative Production Networks

A concept for blockchain utilization in order to schedule resources in collaborative production networks is proposed (see Fig. 3).

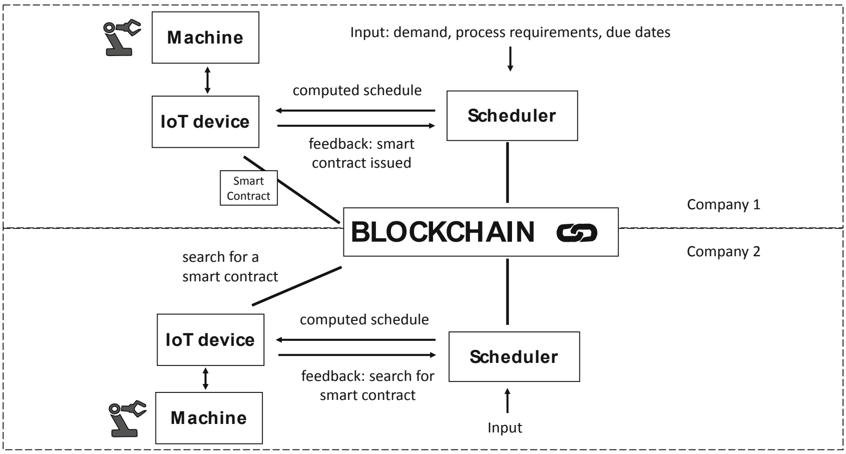


Fig. 3. IoT scheduling process using the blockchain

We consider a collaborative production network with independent network partners. There are several entities involved: manufacturing machines, IoT devices, software, computing nodes and humans. Every entity in the network is represented by a unique address. Each network partner in the network has its own production facilities and operates these irrespective of the network. The main feature of our system is that the P2P network is established between the machines themselves, enabling them to communicate and coordinate directly. Customer demand is accumulated as orders in each company over time. The available machines are scheduled accordingly and the schedule is sent to the machines for processing. Each physical machine is connected to an IoT device, which is a virtual twin of the machine. The IoT device is able to sense the functional state of the machine and process the accumulated data. It is also responsible for the data communication with the scheduling system. There is an IoT device for each process capability that is required in a company in the network. If a time slot on a machine remains unused according to the calculated schedule and is not blocked for maintenance, the machine itself shares the capacity in the network. It issues a smart contract on the blockchain containing information on the available time for processing, processing capabilities and an expected compensation for utilization. The available time is the time between the last scheduled job and the next job to be processed, including setup times. Processing capabilities can either be defined as the setup configuration of the machine or all executable processing types. The calculation basis for the expected compensation has to be decided by the consortium, it could be rather static like machine-hour rates or demand-based. If a company lacks the machine capacity to carry out a process step of an order, the responsible machine as an IoT device automatically searches for an available smart contract with the required key data. If the detected data matches the internal specifications, the process of using the external machine capacity can be triggered directly by a transaction to the address of the smart contract. Transportation can also be initiated by the smart contract. If a smart contract published by a machine is no longer feasible, for example due to a machine defect, the machine can create a new smart contract in the same way as when releasing capacity. The new contract can trigger the previous smart contract through a transaction and thus remove it from the system. Machine capacity is only made available collaboratively in the network if it is not required internally. This ensures that each partner can meet its internal goals and satisfy all stakeholders. Each company retains full control over its machines, only the required data concerning free machine capacities are made transparent. The capacity-providing machines can start processing immediately, as it is ensured that the agreed compensation is received once the transaction has been validated (e.g. by using the underlying cryptocurrency).

The mentioned IoT device is essential for the concept of autonomous intelligent machines in the blockchain scenario. It enables communication across multiple layers in the proposed architecture, as shown in Fig. 4. The physical layer represents the machine itself and is connected to the detection layer that is able to communicate with the data processing layer. The detection layer uses sensor data, adapters and application programming interfaces to provide input and output functions for the physical layer. The related data is processed in the data processing layer in a computing unit. This layer contains drivers for the sensors and adapters, a local storage unit, a control unit for configuring the device for data processing and a blockchain client. The data processing layer also handles communication with the application layer and the blockchain layer. The blockchain client sends and receives transactions to and from the network and transmits smart contracts. The blockchain wallet and the private keys are managed in the client as well. The application layer comprises the application software such as ERP or MES systems, which forward data to the IoT device and receive feedback from the device. In our proposed system, the scheduling software would forward the calculated schedule to the IoT device, which interprets the input and passes it on to the physical machine to start operation. The issuance of a smart contract by the machine in the case of free capacity or the search for a suitable resource in the network in the case of a capacitive shortfall is transmitted back to the appropriate software by the IoT device accordingly. The blockchain layer hosts the blockchain node and enables the P2P connection with other nodes in the network. The application layer could become obsolete in future concepts of Industry 4.0 in which the IoT devices communicate autonomously with the parts to be manufactured.

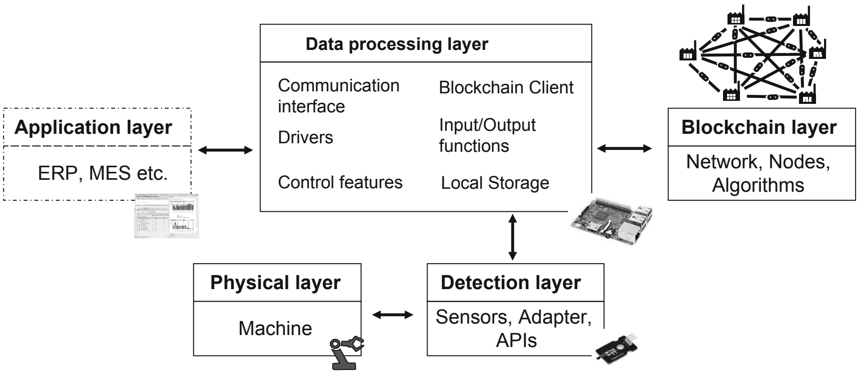


Fig. 4. Architecture of the proposed blockchain system

On the point of consensus building in our proposed system, we suggest a consortium blockchain rather than a public or private blockchain. Once the consortium of n partners has been formed, each partner must operate a full node. The entire blockchain is stored decentrally at n different places. We further propose a Proof-of-Authority (PoA) consensus mechanism, as all network partners are known and presumed to be reputable. Proof-of-Work (PoW) remains useful for networks with a truly minimal level of trust as an alternative solution. The unique identities of all authorities are kept in a smart contract permission list that governs the block creation process. The right to read data from the blockchain could be made public or also be restricted to the consortium in a case-by-case decision.

Our concept goes beyond the existing approaches of blockchain technology in SCM and manufacturing. By indicating a new use case, automatic scheduling of resources in collaborative production networks, the key features visibility and integrity are enhanced (Blossey et al. 2019). The blockchain transparently stores information in a decentralized way and enables a fair consensus mechanism. It can be used to optimally utilize machine capacities in the network in the sense of virtualization or tokenization. In contrast to the concept of cloud manufacturing, production data is not stored centrally or managed by an intermediary in our concept but is available decentrally and transparently. All partners have equal rights in the system and act independently of a cloud service provider.

Although the proposed methodology offers several advantages, there are certainly some general risks involved. The general consideration of joining a collaborative network is accompanied by several opportunities but competitive losses are also possible due to blocked machine capacity. The precise structuring of smart contracts is immensely important and should be carefully assessed, as they are immutable and critical to maintain trust in the network. For networks with very distant partners, the factor of transport time has to be added to the calculations. Legal disputes are partially backed by the transparent history of the blockchain, but cannot be completely ruled out. It is theoretically possible to form a cartel within the consortium blockchain network and influence the election of compromised nodes.

5 Case Study

To demonstrate the conceptual workflow of our proposed system, we use a simple fictional case study with three independent companies (C1–C3) participating in distributed manufacturing in a network. Each company owns two different machine types and one machine per type in a job shop organization. The operations *Op.* for each job, respective machine types *M* and the processing duration for each operation *t* are displayed in Fig. 5. *DD* represents the due date for each job.

C1					C2					C3				
Job	Op.	M	t	DD	Job	Op.	M	t	DD	Job	Op.	M	t	DD
1	1	A	1		1	1	A	5		1	1	B	3	
	2	B	4	8		2	B	3	15		2	A	2	8
2	1	A	2		2	1	B	2		2	1	A	4	
	2	B	5	12		2	A	1	11		2	B	2	8
3	1	A	1		3	1	A	4		3	1	B	1	
	2	B	4	9		2	B	6	10		2	A	6	10

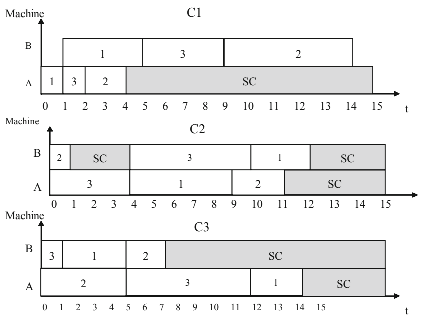


Fig. 5. Job list and Gantt charts

Each company initially schedules its own machines, in this case according to Jackson's algorithm (Jackson 1956). Each free time slot with $t > 1$ is published as a smart contract on the blockchain and denoted as SC in the resulting Gantt charts in Fig. 5. It is evident that C1 cannot complete job 2 on time ($t = 14$, *due date* = 12). The IoT device of machine B (C1-B) starts a search for a smart contract on the blockchain with the requirements time ($4 \leq t \leq 12$), cost ($cij \leq x$) and process capabilities (*machine type* = B). The device detects the smart contract of C3-B accordingly and checks the specified data. As the costs are within reasonable limits, IoT device C1-B automatically triggers the smart contract by sending a transaction, which contains data on the operation to be performed, the due date and the compensation in form of the underlying cryptocurrency. The transaction is included in the next block to be created and validated by the authority nodes. The transaction is now valid and transparent to everyone on the network. The material of job 2 is prepared for transport to C3-B after production on C1-A. This is based on the communication of the IoT device with the blockchain and the application systems. The operation is carried out immediately after the jobs of C3 are finished on C3-B and transported back to C1 or directly to the customer. In this way, C1 is able to complete the order before the due date, increase turnover and satisfy the customer. On top, C1 could provide the machine capacity of its own machine group B to the network by issuing a new smart contract. C3 is able to fully utilize its machine and receives a compensation for the service, which makes an additional contribution to the turnover. The resulting schedule is shown in the Gantt chart in Fig. 6.

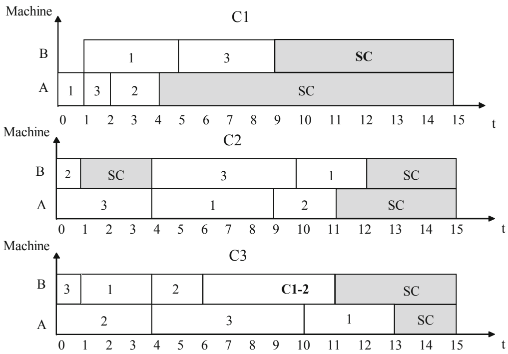


Fig. 6. Optimized schedule using the proposed system for resource sharing

Of the three nodes in the network, two would have to operate as authority nodes. Apart from the cost of setting up the system, transaction costs are low as the PoA consortium blockchain is utilized. The costs mainly depend on the time in which the next block is mined which could be around half a second to utilize the real-time potential of IoT devices in our concept.

6 Limitations, Implementation Challenges and Future Research

Our study holds some limitations that are discussed in this section. First, the concept was not tested in a productive environment. We aim to generate this empirical validation of the concept in future research. There are apparent implementation challenges in real production environments: The choice of a suitable blockchain platform needs to be investigated in adequate detail. Criteria like speed, efficiency in processing transactions, security and ease of use have to be considered. A close collaboration with legal entities is advisable to properly use smart contracts in complex manufacturing environments. From a technical viewpoint, there might be more efficient solutions than a consortium blockchain with as many full nodes as participants. The choice of an appropriate consensus mechanism has to be assessed. Conspiracy among partners in the network could alleviate trust in the technology itself. However, PoA seems to be a sensible choice. Regulation and general governance issues continue to be an issue as the technology evolves. Furthermore, new data protection regulations must be observed. Future research should address the specific legal structure of the network and aim for a proof of concept. It further is advisable to study other areas in SCM and operations that may profit from new digital technologies like the blockchain. A connection of the proposed system with a negotiation approach to optimally allocate demand and supply is also a topic for further research. In-depth case studies can support the empirical validation of the technology's applicability. Master data management in supply chain interaction or the use of cryptocurrencies for reimbursement of manufacturing services appear to be promising research directions.

7 Conclusion

This paper uses the method of design science to present a methodology for implementing a blockchain system with the use of smart contracts in a distributed collaborative production network. To ensure relevance of the constructed methodology and to answer our first research question, we assessed the issues persisting in collaborative production networks and demonstrated how the blockchain contributes to solve these issues. Then we answered our second research question and proposed a methodology for distributed manufacturing using the blockchain. The capabilities of the methodology as well as potential benefits and risks are presented accordingly. A case study with several connected companies using a shared resource is used to demonstrate the applicability of blockchain as a technology to schedule resources in a P2P network with maximized transparency, trust and IoT readiness. The proof of concept is a step for further research as the appropriate environment has to be determined.

References

Abeyratne, S.A., Monfared, R.P.: Blockchain ready manufacturing supply chain using distributed ledger. Int. J. Res. Eng. Technol. **05**, 1–10 (2016)

Angrish, A., Craver, B., Hasan, M., Starly, B.: A case study for blockchain in manufacturing: "FabRec": a prototype for peer-to-peer network of manufacturing nodes. Procedia Manuf. **26**, 1180–1192 (2018)

Bahga, A., Madisetti, V.K.: Blockchain platform for industrial Internet of Things. J. Softw. Eng. Appl. **09**, 533–546 (2016)

Bartsch, F., Neidhardt, N., Nüttgens, M., et al.: Anwendungsszenarien für die Blockchain-Technologie in der Industrie 4.0. HMD Prax der Wirtschaftsinformatik **55**, 1274–1284 (2018)

Benkler, Y.: The Wealth of Networks. Yale University Press, New Haven and London (2006)

Blossey, G., Eisenhardt, J., Hahn, G.J.: Blockchain technology in supply chain management : an application perspective. In: Proceedings of the 52nd Hawaii International Conference on System Sciences, pp. 6885–6893 (2019)

Casino, F., Dasaklis, T.K., Patsakis, C.: A systematic literature review of blockchain-based applications: current status, classification and open issues. Telemat. Inform. **36**, 55–81 (2019)

Das, S.: Blockchain global analysis. In: Cogniz. - Blockchain Glob. Anal. (2017). https://www.cognizant.com/Resources/cognizant-blockchain-global-analysis.pdf. Accessed 10 Jan 2019

evan GmbH: evan.network (2019). https://evan.network/. Accessed 9 Jan 2019

Feng, Q., He, D., Zeadally, S., et al.: A survey on privacy protection in blockchain system. J. Netw. Comput. Appl. **126**, 45–58 (2019)

Hevner, A., Ram, S., March, S.T., Park, J.: Design science in information systems research. MIS Q. **28**, 75–105 (2004)

Hyperledger: Hyperledger – Open Source Blockchain Technologies (2018). https://www.hyperledger.org/. Accessed 20 Dec 2018

Jackson, R.: An extension of Johnson's results on job lot scheduling. Nav. Res. Logist. Q. **3**, 201–203 (1956)

Kagermann, P.D.H., Wahlster, P.D.W., Helbig, D.J., et al.: Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0 (2013). https://www.bmbf.de/files/Umsetzungsempfehlungen_Industrie4_0.pdf. Accessed 20 Dec 2018

Kim, H., Laskowski, M.: A perspective on blockchain and smart contracts. In: IEEE Work Work Privacy, Secur Trust Blockchain Technol 6 (2017)

Kohtala, C.: Addressing sustainability in research on distributed production: an integrated literature review. J. Clean. Prod. **106**, 654–668 (2015)

Korpela, K., Hallikas, J., Dahlberg, T.: Digital supply chain transformation toward blockchain integration. In: Proceedings of the 50th Hawaii International Conference on System Sciences, pp. 4182–4191 (2017)

KPMG: Global manufacturing outlook - competing for growth: how to be a growth leader in industrial manufacturing (2016). https://home.kpmg/content/dam/kpmg/pdf/2016/05/global-manufacturing-outlook-competing-for-growth.pdf. Accessed 8 Jan 2019

Li, Z., Barenji, A.V., Huang, G.Q.: Toward a blockchain cloud manufacturing system as a peer to peer distributed network platform. Robot. Comput. Integr. Manuf. **54**, 133–144 (2018)

Lou, P., Ong, S.K., Nee, A.Y.C.: Agent-based distributed scheduling for virtual job shops. Int. J. Prod. Res. **48**, 3889–3910 (2010)

Makhdoom, I., Abolhasan, M., Abbas, H., Ni, W.: Blockchain's adoption in IoT: the challenges, and a way forward. J. Netw. Comput. Appl. **125**, 251–279 (2019)

Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008). https://bitcoin.org/bitcoin.pdf. Accessed 10 Jan 2019

Pai, S., Sevilla, M., Jerome, B., et al.: Does blockchain hold the key to a new age of supply chain transparency and trust? (2018). https://www.capgemini.com/wp-content/uploads/2018/10/Digital-Blockchain-in-Supply-Chain-Report.pdf. Accessed 20 Dec 2018

Parity: Parity technologies: proof-of-authority chains - parity tech documentation (2019). https://wiki.parity.io/Proof-of-Authority-Chains. Accessed 28 Jan 2019

Petersen, M., Hackius, N., von See, B.: Mapping the sea of opportunities: blockchain in supply chain and logistics. IT Inf. Technol. **0**, 263–271 (2018)

Pinedo, M.: Scheduling - Theory, Algorithms, and Systems, 4th edn. Springer, New York (2012)

Ren, L., Zhang, L., Wang, L., et al.: Cloud manufacturing: key characteristics and applications. Int. J. Comput. Integr. Manuf. **30**, 501–515 (2017)

Renna, P., Argoneto, P.: Capacity sharing in a network of independent factories: a cooperative game theory approach. Robot. Comput. Integr. Manuf. **27**, 405–417 (2011)

Reyna, A., Martín, C., Chen, J., et al.: On blockchain and its integration with IoT. Challenges and opportunities. Futur. Gener. Comput. Syst. **88**, 173–190 (2018)
Rodríguez Monroy, C., Vilana Arto, J.R.: Analysis of global manufacturing virtual networks in the aeronautical industry. Int. J. Prod. Econ. **126**, 314–323 (2010)
Rossit, D.A., Tohmé, F., Frutos, M.: Industry 4.0: smart scheduling. Int. J. Prod. Res. 1–12 (2018)
Rupareliya, P.: How can Blockchain-as-a-Service fuel a new phase of innovation among global enterprises? (2018). https://medium.com/swlh/blockchain-as-a-service-platforms-5f8c97850d92. Accessed 16 Jan 2019
Scott, T.: TradeLens (2018). https://www.ibm.com/blogs/think/2018/11/tradelens-how-ibm-and-maersk-are-sharing-blockchain-to-build-a-global-trade-platform/. Accessed 16 Jan 2019
Shi, Y., Fleet, D., Gregory, M.: Global manufacturing virtual network and its position in manufacturing systems. In: The 7th Annual International Manufacturing Symposium, pp. 197–205 (2005)
Srai, J.S., Kumar, M., Graham, G., et al.: Distributed manufacturing: scope, challenges and opportunities. Int. J. Prod. Res. **54**, 6917–6935 (2016)
Staufen, A.G.: Deutscher Industrie 4.0 Index 2018 (2018). https://www.staufen.ag/fileadmin/HQ/02-Company/05-Media/2-Studies/STAUFEN.-Studie-Industrie-4.0-Index-2018-Web-DE-de.pdf. Accessed 20 Dec 2018
Szabo, N.: Formalizing and securing relationships on public networks (1997). http://firstmonday.org/ojs/index.php/fm/article/view/548/469. Accessed 9 Jan 2019
Tao, F., Cheng, Y., Zhang, L., Nee, A.Y.C.: Advanced manufacturing systems: socialization characteristics and trends. J. Intell. Manuf. **28**, 1079–1094 (2017)
Tuma, A.: Configuration and coordination of virtual production networks. Int. J. Prod. Econ. **56–57**, 641–648 (1998)
Wang, L., Shen, X., Li, J., et al.: Cryptographic primitives in blockchains. J. Netw. Comput. Appl. **127**, 43–58 (2018)
Yang, X., Shi, G., Zhang, Z.: Collaboration of large equipment complete service under cloud manufacturing mode. Int. J. Prod. Res. **52**, 326–336 (2014)
Zheng, Z., Xie, S., Dai, H.-N., et al.: Blockchain challenges and opportunities: a survey. Int. J. Web Grid Serv. **14**, 1–25 (2018)
Zheng, Z., Xie, S., Dai, H., et al.: An overview of blockchain technology: architecture, consensus, and future trends. In: Proceedings of the 2017 IEEE 6th International Congress on Big Data, BigData Congress 2017, pp. 557–564 (2017)

Realizing the Full Potential of Robotic Process Automation Through a Combination with BPM

Christian Flechsig(✉), Jacob Lohmer, and Rainer Lasch

Chair of Business Management, esp. Logistics,
Technische Universität Dresden, 01062 Dresden, Germany
{christian.flechsig, jacob.lohmer,
rainer.lasch}@tu-dresden.de

Abstract. Robotic Process Automation (RPA) is currently a topic of interest for both research and industry. Many administrative tasks in operations consume a vast amount of time and create little value. The possibility that the responsible employees can quickly and easily automate these processes by themselves makes RPA a promising approach for many processes. However, RPA procedures are only able to automate a process in its present form. In this way, redundancies and excessive process steps are incorporated into RPA process flows. Combining RPA with the popular Business Process Management (BPM) approach poses a useful strategy as the *as-is* process is optimized first. The RPA procedure is then deployed on the optimized process to reach its full automation potential. This paper proposes a methodology to combine RPA and BPM and demonstrates the potential within a case study. Benefits, limitations, and further research opportunities are likewise addressed.

Keywords: RPA · BPM · Automation · Combination

1 Introduction

To remain competitive in today´s complex and competitive economic environment, process automation is essential for companies of all sizes [1]. Efforts for digitalization and process automation in production environments have been increased considerably in the last years in the context of Industry 4.0. Several manufacturing processes on the shop floor have supplied a vital base for optimization and digitalization that have enabled companies to efficiently produce complex products at low costs. These efforts have led to an increase in administrative work tasks, which are often repetitive, time-consuming and have not been examined with the same care and accuracy as manufacturing process automation has received [2, 3]. One of the reasons for this discrepancy is that the output of a business process cannot be measured simply on the basis of quality and quantity like a manufacturing process, since the process result is usually more complex in business processes. Science provides general methods to automate business processes to reduce administrative efforts in line with concepts like Lean Administration [4, 5].

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 104–119, 2019.
https://doi.org/10.1007/978-3-030-29821-0_8

Robotic Process Automation (RPA), a comparatively new method, poses the advantage that the responsible employees themselves can quickly and easily automate these administrative processes. However, RPA procedures are intended to automate a process at hand without an in-depth analysis or valuation of the necessity and usefulness of automated process steps [6]. As a consequence, redundancies and dispensable process steps are also transferred into new designed RPA process flows. Incorporating the procedures of RPA into the well-known Business Process Management (BPM) approach appears to be a useful strategy. The *as-is* process could be standardized and optimized first. Consequently, the RPA procedure would be able to reach its full automation potential on the remaining process steps. As the combination of BPM and RPA has rarely been addressed in scientific literature, this paper intends to propose a methodology and demonstrate the potential of the combined use of BPM and RPA within a case study. Specifically, we intend to answer the following research questions:

RQ1: How can BPM and RPA be distinguished?
RQ2: Is there synergy potential between BPM and RPA? If so, how could a combination of both methods be achieved?
RQ3: What are the benefits and risks of this concept?

The remainder of this paper is organized as follows: In the next section, we provide the background of our study with an overview of the characteristics of BPM and RPA. The literature on both concepts is reviewed. We compare BPM and RPA and demonstrate the synergy potential. In the third section, we present a methodology to combine RPA and BPM in digitalization projects. Afterwards, we show the potential of our proposed methodology in a case study and discuss capabilities, benefits, and risks involved. To conclude, the last section presents limitations, further research opportunities, and an outlook.

2 Comparison of BPM and RPA and Synergy Potential

Different concepts can be utilized for the automation of business processes. This section provides an overview of the specific characteristics of the two concepts of interest for this study: BPM and RPA. First, both are defined and compared. Then, the synergy potential between BPM and RPA is evaluated and indicated.

2.1 Characteristics of BPM and RPA

In order to clearly distinguish between BPM and RPA and to reveal the benefits of a combined application, both methods are defined in this section, including the elaboration of differences and similarities.

BPM. In scientific literature, BPM is a vague term that lacks a precise definition and delimitation [7]. The *European Association of Business Process Management* [8] specifies BPM as follows: “Business Process Management (BPM) is a systemic approach geared to capture, design, execute, document, measure, monitor and control automated as well as non-automated processes in order to meet the objectives that are aligned with the business strategy of a company. BPM embraces the conscious, comprehensive and increasingly technology-enabled definition, improvement, innovation and maintenance of end-to-end processes." By means of BPM, the process is understood as a whole and optimized from an organizational perspective with the help of information technology (IT). The conceptual basis of BPM has been developed in the 1990s: Business Process Reengineering (BPR) led to a radical redesign of business processes as they were increasingly recognized as drivers for efficiency and innovation [7, 13]. Business processes can be divided into sub-processes and mainly consist of administrative activities, which are planned, controlled, and executed in a predefined order to achieve a specific goal. They represent procedures within or across the company that require time, resources, and expenses and are triggered by internal as well as external events [9–11]. Through BPM, processes can be aligned with the business strategy, improving the overall performance due to process optimization within single departments, between multiple departments of the company (inter-divisional) or even beyond company borders [8]. Further advantages are reduced costs and increased quality of processes, organizational flexibility and performance as well as customer and employee satisfaction [12].

The automation of the considered business processes can be realized with Business Process Management Systems (BPMS). These IT systems allow for a precise definition of processes by means of one or several process model notations to analyze, model, and simulate them. Additionally, they facilitate the executional support and integration into other systems [11]. The present de-facto standard, Business Process Model and Notation (BPMN), also offers process execution methods [14].

RPA. Robotic Process Automation (RPA) is a general heading for software tools operating on the user interface of IT systems. Thereby, the term "robot" technically describes a single software license [15]. RPA flows imitate human behavior and relieve or replace employees by automating simple, repetitive, and rule-based back office business tasks [16]. Studies suggest that RPA software robots could lead to cost savings ranging between 20% and 50% compared with a full-time employee while performing the same work amount [6, 17]. Further benefits of RPA are good scalability, enhanced process accuracy, and transparency. Software robots work around the clock and can be trained by refining and extending the script [18, 19].

However, BPM and RPA differ in essential criteria. The following comparison (see Table 1 for a comprehensive overview) focuses on the most relevant distinguishing characteristics used in literature: area of application, procedure of automation, method of integration in existing IT systems, required personnel for implementation, estimated implementation effort, scope of the introduction phase, and dissemination in practice and academia.

Area of Application. BPM is a holistic management concept facilitating the optimal redesign of all kind of processes in an enterprise. Concerning process automation, the primary area of BPMS are rather complex, extensive, and specialized processes with a high proportion of added value and strategic significance across the whole company. Whereas RPA automates simple, already existing repetitive processes with little strategic impact and predominantly local effects on the respective department [6].

Table 1. Characteristics of BPM and RPA

Criterion	BPM	RPA
Area of application	Redesign of extensive processes with high strategic relevance and added value	Automation of simple, already existing repetitive processes with lower strategic relevance
Procedure of automation	Analysis (as-is model), redesign (to-be model), test, implementation, monitoring and control	Technical analysis (as-is model), technical design, test, implementation
Method of integration	"Top-down" via API and web services (logic layer), modification of IT systems necessary	"Bottom-up" via GUIs of existing systems (presentation layer), IT systems remain unchanged
Implementation personnel	Business analysts (model) and specialized software developers (code)	Qualified Business analysts (model and code) or specialists
Implementation effort	High costs and long amortization (programming skills required)	Low costs, quick amortization (less programming skills necessary)
Introduction phase	Extensive system test phase, continuous monitoring	Short system test phase, potentially high training expenditure
Dissemination	Widespread, advanced knowledge	Limited in research, great interest from practice

The technology is also able to make simple decisions following a precisely defined set of criteria. Typical tasks of RPA flows include searching, analyzing, transforming, and saving data in various software applications [18].

Procedure of Automation. The sequence of activities in a BPM project can be described in life cycle models, of which various scientifically illustrated versions exist (for an overview of different models, see [9]). We focus on the model of [11], since it advances existing models (see Fig. 1). In the first phase, relevant processes are identified and defined precisely. Then, the current state of each process is documented, also called *as-is* modeling. The resulting process models reflect and facilitate the understanding of the responsible employees. Stakeholder communication can be improved using the *as-is* models. Their quality has a significant impact on the successful realization of BPM [20]. The following phase (process analysis) identifies and quantifies issues associated with the *as-is* models to reveal optimization potential. The processes can then be purposefully redesigned and improved by defining appropriate alternatives, also called *to-be* modeling. Next, the resulting *to-be* processes with its necessary changes are put into practice. The realization covers two aspects: organizational change management, which includes modifications concerning the way of working, and process automation which refers to the implementation of suitable IT systems. Once the redesigned business process is running, continuous monitoring and analysis of data concludes the BPM life cycle. This data collection can be performed through process mining that supports a targeted process analysis and monitoring [21].

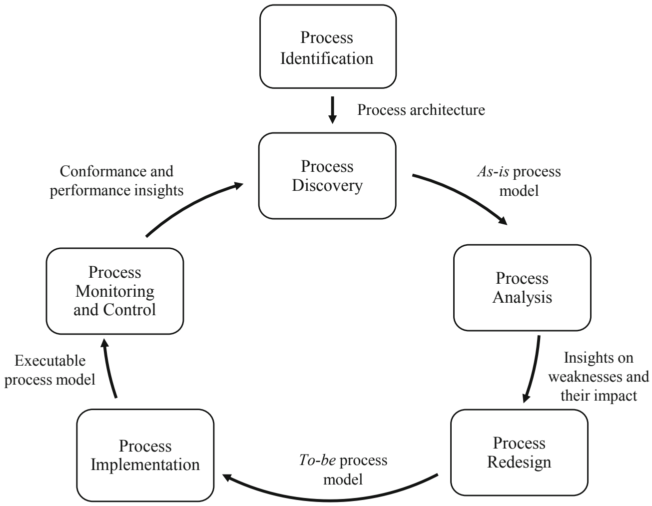


Fig. 1. BPM life cycle model by [11]

RPA does not provide an optimized *to-be* model. The initial *as-is* process is rather analyzed regarding its technical realization and less concerning its optimizing potential. It is then mapped in the RPA tool language as a detailed and executable script. There are two ways of creating the source code: screen scraping and hard coding. Using the first method, the RPA software is recording all user entries and transferring them to source code automatically. By means of hard coding, the data is directly embedded into the source code [6, 22]. After a successful testing phase, the RPA flow is implemented in the productive environment performing "if, then, else" sequences. However, a holistic process monitoring and controlling is missing [23, 24].

Method of Integration. BPMS use various application programming interfaces (API) or web services to access third-party systems and improve them. They generate new models and databases in the data layer. Integration takes place "top-down", since processes are standardized through implementation. RPA programs are directly based on the graphical user interface (GUI) of existing IT systems. They are integrated into the presentation layer and do not change the deeper system logic. Integration takes place via direct process integration prior to implementation ("bottom-up") [15–17].

Required Personnel for Implementation. Process modeling in BPM is predominantly performed by business analysts. However, specialized software developers are needed for the implementation in the BPMS. RPA is almost completely independent of IT staff. Programming can be carried out directly in the respective departments after an initial configuration phase [25, 26].

Implementation Effort. The implementation effort for BPM is significantly higher than for RPA. Full-scale BPM projects have significantly higher costs and payback periods due to the additional technical effort involved [6, 15, 16].

Introduction Phase. Process automation with BPM requires an extensive test phase as well as continuous monitoring of the system. RPA, on the other hand, profits from simple implementation and a short testing phase. The training duration of the software robots certainly increases with faulty programming and insufficient definitions of process sequences [26].

Dissemination. Due to its importance for both academia and practice, BPM constitutes an established discipline and is considered as the most comprehensive and well-known business process management approach [27]. BPM is used widely in practice and shows an advanced state of knowledge in the scientific literature [26]. The modeling language BPMN has been the subject of many recent research projects. RPA has so far shown a rather low dissemination, although practice is familiar with its potential benefits and is beginning to focus more on this area. However, scientific literature on RPA is limited [17]. Relevant contributions mainly present case studies [28]. The combination of RPA with artificial intelligence to automate creative processes is a field that has recently moved into focus [29]. Intelligent RPA (IRPA) recognizes unstructured data such as e-mails or handwritten recordings using techniques like natural speech analysis. In the event of an unknown situation, the IRPA system acts independently and develops continuously [28]. This enables the software robot to react more flexibly to changing tasks. Although intelligent automation is promising, it is still at an early stage of development and is rarely studied scientifically. Moreover, IRPA requires higher expenses and a longer implementation time than standard RPA [29].

Besides their differences, BPM and RPA show several similarities. Both concepts enable the systematic automation of business processes [6]. They are applied for consistent IT support and therefore use similar components, such as graphical editors for modeling [18]. Common goals are the exploitation of potential, e.g. by reducing process costs and time as well as increasing customer satisfaction and frequency [11, 16]. The organizational change associated with the respective projects has to involve all stakeholders at an early stage [12, 25].

The above comparison contributes to answering the first and part of the second research question. BPM and RPA can be clearly distinguished; they have reciprocal advantages and disadvantages. In practice, they do not substitute each other, but are used simultaneously and context-dependent [6]. We provide indications to utilize the synergy potential in the following, leading to an integration of RPA into a holistic BPM concept.

2.2 Synergy Potential Between BPM and RPA

A combination of both methods provides three major advantages. Firstly, it would benefit RPA as this method is not suitable for extensive processes including multiple process steps in principle. However, RPA could be possibly applied if the process is simplified and optimized in advance through a BPM *to-be* model. Secondly, the RPA flow would benefit from the holistic BPM monitoring and controlling, which comprises all business processes that are executed in the company, enabling a faster detection of further optimization potential. And thirdly, BPM benefits from a proper combination as RPA enhances the possibilities to automate processes that are economically not feasible regarding a costly BPM project. In consequence, the holistic BPM approach obtains a further automation concept in addition to BPMS to automate more processes as shown in the following.

Figure 2 depicts the positioning of BPM and RPA in terms of feasibility and process characteristics. The x-axis shows the different types of cases (processes). Similar cases that are managed in the same way are assigned to the same case type. The y-axis shows the frequency of case types (number of similar cases) in a given period. The frequency needs to be individually determined for each company. The correlation of both axes is typically described by a Pareto distribution. A lot of case types are quite infrequent; 80% of all cases can be explained by 20% of the case types. Traditional automation through BPMS focuses on the most frequent 20% of case types (dashed line in the left sector). These processes are often highly value-adding with many similar process steps and only a few exceptions, making automation economically feasible. The remaining 20% of cases contain 80% of the case types and are often performed manually by humans (solid line in the middle and dotted line in the right sector). This includes repetitive work and decision making due to exceptions and various process steps. These types of cases are more time-consuming than the frequent ones. For a small proportion of the processes (dotted line), automation is not possible or economically feasible. However, a large proportion (solid line) can be automated with RPA [15]. Such processes are usually of low to medium complexity and can be performed within 5 to 30 min [17]. RPA supports employees by executing tasks that are repetitive but not frequent enough to justify traditional automation through costly BPM projects. Therefore, RPA constitutes a sensible alternative as it is applied using the existing GUIs instead of adjusting to changing APIs. Moreover, it could save time and resources to automate suitable processes with RPA instead of BPM, resulting in shorter amortization cycles and test periods. Together, a combination of BPM and RPA would therefore cover most if not all processes that can be automated to some extent (dashed line and solid line).

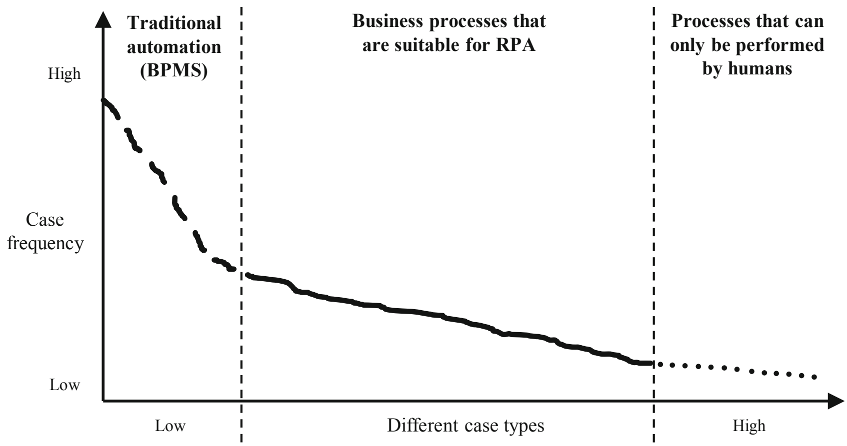


Fig. 2. Positioning of BPM and RPA (based on [15])

Furthermore, by means of a combined BPM-RPA methodology, the BPM system can trigger an RPA activity to interact with other IT systems and use the returned results in the subsequent steps of the BPM process flow. Other promising process types are those that require human approval after the transfer of the result or due to necessary exception handling. BPM can initiate these processes based on the results of an RPA flow. This minimizes potential human error while allowing human control over the outcome. The results of the approval process could be logged for compliance requirements [30]. Our proposed approach is presented in the following.

3 Methodology to Combine BPM and RPA

We developed a methodology to combine BPM and RPA. Figure 3 demonstrates our proposed life cycle model. The first four phases are inspired by the existing BPM life cycle of [11] (see Fig. 1). The *as-is* model is established, analyzed, and optimized to create a *to-be* process model. After the redesign and optimization, a case distinction is necessary. Based on Sect. 2.2 and Fig. 2, the users need to decide whether the process under consideration is suitable for RPA or not. This decision can be supported by several questions: Is the process repetitive and standardized (rule-based)? Is the error potential high? Is permanent processing necessary? Is the process relevant for compliance? Does the process occur rather frequently and at least partially contribute to value creation? Are resources such as budget or IT staff currently scarce? If all or most of these questions are answered affirmatively, the process appears suitable for an automation with RPA. Furthermore, the process should require neither subjective judgement nor creativity but should be clearly defined and highly standardized.

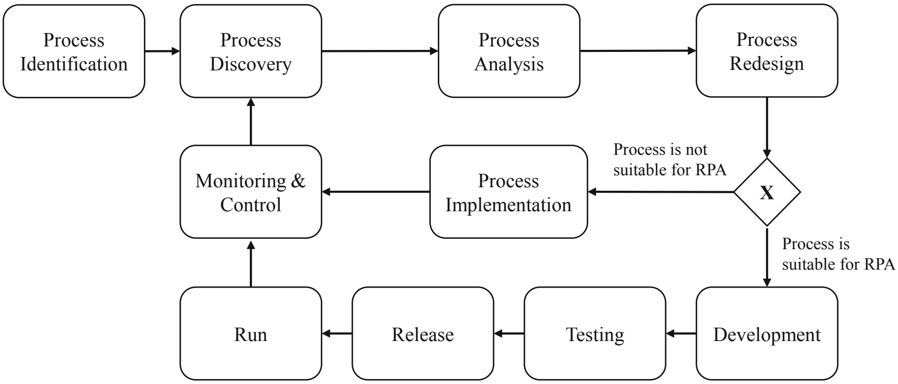


Fig. 3. Proposed BPM-RPA life cycle model

Processes that are suitable for RPA then deviate in the life cycle model and pass through the development, testing, release, and run phases [23]. In the development phase, a technical design and an implementation plan are created. The automation procedure follows the standardized *to-be* process model. Testing ensures the usability of the RPA flow, especially the consideration of all specifications and a proper robot functionality. Testing should be carried out in a non-productive environment with real, historical data. The release phase represents the transition from the test to the productive environment, which is planned and then executed. After the release, the process is performed for several iterations while being constantly monitored by a human employee to ensure stable performance in live operations and the achievement of expected benefits. As soon as the correct sequence has been guaranteed, the BPM flow is reactivated and takes over monitoring and control accordingly. The processes that are not suitable for RPA follow the known path with process implementation as the next step and meet again with the RPA flow in the step monitoring and control.

Our life cycle model provides the following benefits: The BPM process is not profoundly changed, but the RPA flow is added at a critical point. The decision for RPA process suitability is determined after the target process modeling has been completed. This guarantees full exploitation of the optimization potential of the BPM procedures. Although the effort involved will be greater if an upstream optimization is performed, we avoid a premature execution of the RPA sequence based on the *as-is* process. After the successful implementation of an RPA flow, this sequence is also subject to monitoring and control, enabling the same control level as for the remainder of the BPM process. Iteratively, the process is checked and optimized again. RPA can thus be replaced by a full automation of the affected IT systems at a later point in time, when missing resources such as budget or IT staff become available. Anyway, if a (costly) BPM project starts, RPA suitability should be checked to prevent unnecessary costs and efforts. With our combined life cycle model, more processes can be covered (see dashed line and solid line in Fig. 2) and the methods benefit from the advantages of each other. To demonstrate the potential of our proposed methodology, we present a case study hereafter.

4 Case Study

The case study is inspired by a typical process of our case company and can be attributed to production planning and control. The case company operates in the semiconductor industry, with a high level of automation in the manufacturing process. However, many repetitive planning processes are still performed manually at high costs.

Production planners solve scheduling tasks daily that are handled differently, depending on the assigned process flow of each order. Our case company deals with a large product mix and many manufacturing steps. Thus, it uses a wide variety of IT systems and databases to manage and change process data. Particularly in industries such as microelectronics, product life cycles have become very short and new products are enhanced in productive operations as development batches. Often the process flows are modified iteratively and are not immediately available in the operations department, but have to be collected from various databases. Due to its generality, our case study can also be adopted for companies operating in other industry sectors, showing comparable results.

The *as-is* process of the case study contains the following process steps: The order list is received in the operations department as an MS Excel worksheet. It consists of six columns: a serial number (ID), information about the product ordered and its demand, as well as its production steps with related characteristics (see Table 2). The order list contains 300 different products, the product mix consists of 15 product classes with 20 variants. Order quantities range from single orders up to 200 units per order. In our case study, we consider a production scenario with three production steps (X, Y, Z) in different sections (A, B, and C). Each production step can be performed in six different modes. For every production step, the research and development (R&D) department maintains a respective database in which the process characteristic for the product is updated first. Some of the products show missing information regarding one or several production step(s). We created three cases to investigate the impact of the quantity of missing information on the process execution (both manual and RPA). In **case 1** we assume that 5% of the products (n = 15) are affected by missing information. The total amount of lacking data fields is 21. In **case 2**, 10% of the products (n = 30) show missing information, comprising 42 lacking data cells. This is the standard case, based on expert feedback from the case company. Furthermore, we created **case 3** with 20% of affected products (60) and 84 data fields missing. The order list is initially and manually checked by the operations employee for completeness with regard to the production steps. If an entry is missing, the planner checks the information in the respective database that is maintained for the production step. The planner copies/notes the required product type, opens the database, searches for the entry of the product type and for the missing information. If the information is accessible in the database, the employee copies the data and pastes it into the order list. If the information is missing, the planner sends an e-mail to the responsible colleague in the R&D department to ask for feedback on the specific production step. The affected order is restrained as long as the process flow is not clarified. Once all orders have been checked, the planner releases the orders that have not been restrained.

Table 2. Order list (excerpt)

ID	Product	Demand (units)	Production step A	Production step B	Production step C
1	A1	116	X2	Y4	Z3
2	A2	145		Y4	Z6
…	…	…	…	…	…
299	O19	153	X3	Y5	Z1
300	O20	171	X4		

The previously described *as-is* process was analyzed. Several issues were identified and the process was redesigned. Figure 4 shows an excerpt from the optimized manual *to-be* process model where the three databases were merged into one master database to reduce individual search efforts as well as the complexity of the IT infrastructure. Moreover, it is methodically easy to implement for the IT department. As an additional step, the planner now saves the queried information in the operation department's own database to prevent redundant searches in upcoming weeks. Thus, stored data is automatically used to generate and complete future order lists. The *to-be* process model was subsequently reviewed in order to make a decision on the potential suitability for RPA. The process flow is repetitive, moderately complex, and fairly value adding. It seems reasonable to use an RPA flow for these process steps.

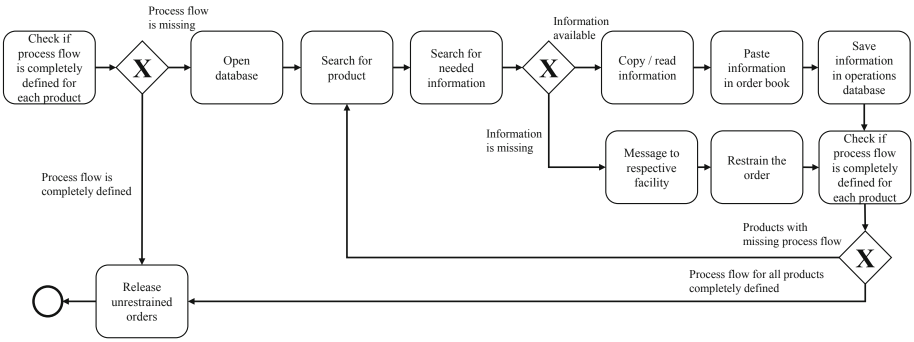


Fig. 4. Manual *to-be* process model (excerpt)

The *to-be* process flow following our methodology with a combined use of BPM and RPA is described in the following. After determining its suitability for RPA, we transformed the manual *to-be* process model (Fig. 4) into a technical RPA-design. This implementation plan must be encompassing and rule-based, defining responsibilities and containing all process steps and a detailed exception handling. Due to its complexity and space limitations, the RPA process model is not included as a figure in this paper. However, we would like to focus on two points in particular: the retrieval of information and the handling of exceptions. The RPA flow automatically searches the master database for missing information detected in the order list. The data is pasted into the order list and saved in the same way as in the manual *to-be* model (Fig. 4). If information is missing in the retrieval database, the RPA flow automatically prepares a list for the operation employee with the respective products and the missing information. A handover notice is created to return control to the human employee. The employee can then easily request feedback from responsible colleagues in the R&D department. If no information is received after a certain period of time (e.g. 30 min), the order is restrained. Any received information is pasted into the order list and the operations departments database accordingly. This procedure ensures the correct handling of each order and enables more orders to be produced in the next period. As the RPA flow is much faster than manual processing, more time can be spent on processing the special cases and the employee can systematically consult with colleagues.

In order to evaluate the practical benefits of our proposed concept, we compare the different workflows and respective cases regarding the time required to process the order list. The workflows are the manual *as-is* model, the manual *to-be* model, the RPA *as-is* model and the RPA *to-be* model. In the *as-is* models, an MS Excel order list and three separate MS Excel databases were used, which were combined into one MS Excel master database for the *to-be* models (MS Office 2016, Windows 10, CPU i5-6500). Both RPA flows were designed and implemented using UiPath Studio v2019.1.0 [31] on the same workstation. The *to-be* process starts with opening the order list and checking for completeness. The master database is then opened (manual *to-be* process) or the RPA flow is initiated by starting the RPA software and the flow execution. The resulting completed order list is finally revised. Figures 5 and 6 show the different runs for the single workflows and cases. The data were obtained over several runs (n = 6) in the case company by observing the corresponding employees (case 2) and by experiments (case 1 and 3).

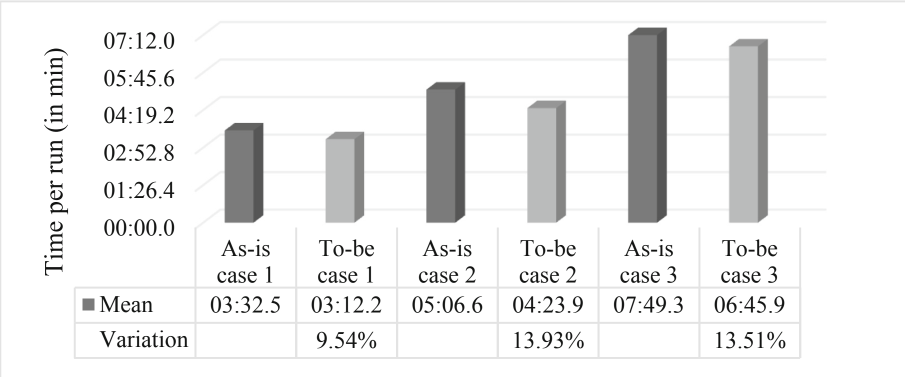


Fig. 5. Comparison as-is, to-be process model

As a result of the optimization (manual *to-be* model compared with the manual *as-is* model), the mean time decreased by 9.54% in case 1, 13.93% in case 2 and 13.51% in case 3. Figure 6 shows the runtimes of the RPA flow for the three cases and the as-is and to-be models. The mean time decreased by 3.49% in case 1 if the RPA flow is executed based on the *to-be* model rather than the *as-is* model. For case 2 and case 3 the results are a decrease of 3.68% and 3.11%, respectively. Overall the runtime of case 1 was improved by 70.53%, case 2 by 79.53% and case 3 by 86.06%. Considering the amount of missing information (cases), the RPA process is insensitive in terms of total runtime. From case 1 to case 2 (twice as many cases), the runtime increases by less than one second, from case 2 to case 3 by less than two seconds.

These results underline the simple but effective optimization from the *as-is* to the *to-be* process model and the high performance of an RPA flow. Due to the short processing times of an RPA application and the comparatively simple and generic case process, the differences between the as-is RPA flow and the optimized to-be RPA flow are rather small. However, the advantage of an RPA execution based on the to-be model instead of the as-is process was shown. The use of the RPA flow also contributes to error prevention. As the automation using UiPath is based on ready-to-use modules in the software and is relatively self-explanatory, a duration of about three working days for an employee from the specialist department or about two hours of

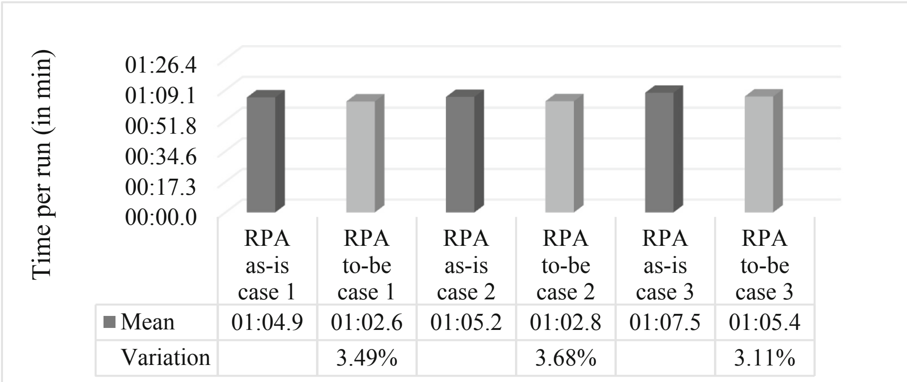


Fig. 6. Comparison of RPA *as-is* and RPA *to-be* model in the three cases

programming effort for a skilled coder can be estimated. Due to the savings in working time in the specialist department, the amortization period is thus around two months if the same wage level is assumed. To summarize the benefits, the new *to-be* RPA flow of our combined approach enables a cost-saving, fast, and error-avoiding process flow.

At a later point in time, our case process could be streamlined even more through a master database that is maintained by the R&D department exclusively. The order list would then only include the demand and the corresponding order quantities for each product. When scheduling, the order list is completed by automatically fetching process information from the R&D database. Thus, the RPA flow could be replaced by a full-scale BPM project that continuously checks the consistency and completeness of the R&D database and inserts the information in the order list.

5 Conclusion

In this paper, we presented a methodology for combining BPM and RPA. After a literature-based comparison to define both concepts and identify distinguishing characteristics, we showed synergy potential of a combination, such as error handling through BPM and the release of RPA results in the BPM process. The optimization of the *as-is* process model in the proposed methodology supports the realization of the full potential of RPA and avoids a premature execution of the RPA sequence. After its successful implementation, this sequence is also subject to the holistic BPM monitoring and control. Thus, the RPA flow can be replaced by full-scale automation later. Our combined approach encompasses more processes than the concepts can individually. Thus, BPM is enabled to automate further activities. The potential of our methodology was shown in the presented case study.

In addition to the mentioned benefits, our study is subjected to some limitations. The proposed approach is a theoretical concept which has not yet been validated in practice. It is based on the BPM life cycle of [11], which in our opinion is the most relevant and significant BPM model among the investigated literature. Other concepts might be more fitting in certain cases. We have automated the given workflow with the software tool UiPath [31]. There are various ways that may lead to slightly different results. Automation through BPMS could result in a faster process, but implementation would be more complex and expensive. A practical implementation considering a real use case could validate the benefits of our approach and show possibilities of useful integration. As RPA promises a significant improvement and can be performed quickly following a clear set of rules in our concept, it could lead to blind faith in this technology. RPA requires clearly defined and appropriate processes to function properly. The relatively simple implementation of RPA could encourage companies to automate unnecessary and inappropriate workflows. Thus, it could hinder critical scrutinizing and further development of significant business processes.

To conclude, BPM and RPA can be clearly distinguished. Both concepts have reciprocal advantages and disadvantages. In practice, they do not substitute each other, but are used simultaneously and context-dependent. Companies should therefore consider integrating RPA in their holistic BPM concept. Since RPA lacks scientific literature, there is a lot of potential for further research. IRPA seems worth investigating, as it could potentially replace traditional BPMS in the future. However, a human optimization from the *as-is* to the *to-be* process model remains necessary. Artificial intelligence is not yet capable of taking this important step on its own. Our methodology could also profit from broader empirical validation in a productive environment, considering more comprehensive processes and an enhanced set of criteria for the RPA usage, which we will aim for in future research.

Acknowledgments. This research was supported by the EU project iDev40. The project iDev40 has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 783163. The JU receives support from the European Union's Horizon 2020 research and innovation program. It is co-funded by the consortium members, as well as grants from Austria, Germany, Belgium, Italy, Spain, and Romania.

References

1. Pourmirza, S., Peters, S., Dijkman, R., Grefen, P.: A systematic literature review on the architecture of business process management systems. Inf. Syst. **66**, 43–58 (2017)
2. Strykowski, S., Wojciechowski, R.: Composable modeling and execution of administrative procedures. In: Kö, A., Leitner, C., Leitold, H., Prosser, A. (eds.) Advancing Democracy, Government and Governance. Lecture Notes in Computer Science, pp. 52–66. Springer, Heidelberg (2012)
3. Alexander, S., Haisermann, A., Schabicki, T., Frank, S.: Robotic Process Automation (RPA) im rechnungswesen und controlling - welche Chancen ergeben sich? Controlling **30**, 11–19 (2018)
4. Brenner, J.: Lean Administration. Carl Hanser Verlag, München (2018)
5. Künzel, H.: Erfolgsfaktor Lean Management 2.0. Springer, Heidelberg (2016)
6. Lacity, M.C., Willcocks, L.P., Craig, A.: Robotic process automation at Telefónica O2. MIS Q. Exec. **15**, 21–35 (2016)
7. Klun, M., Trkman, P.: Business process management – at the crossroads. Bus. Process. Manag. J. **24**, 786–813 (2018)

8. Benedict, T., Bilodeau, N., Vitkus, P., et al.: BPM CBOK Version 3.0: Guide to the Business Process Management Common Body of Knowledge. ABPMP (2013)
9. Alotaibi, Y., Liu, F.: Survey of business process management: challenges and solutions. Enterp. Inf. Syst. **11**, 1119–1153 (2017)
10. Chinosi, M., Trombetta, A.: BPMN: an introduction to the standard. Comput. Stand. Interfaces **34**, 124–134 (2012)
11. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management. Springer, Heidelberg (2013)
12. Haddad, C., Ayala, D., Uriona Maldonado, M., et al.: Process improvement for professionalizing non-profit organizations: BPM approach. Bus. Process. Manag. J. **22**, 634–658 (2016)
13. Margherita, A.: Business process management system and activities. Bus. Process. Manag. J. **20**, 642–662 (2014)
14. Jošt, G., Huber, J., Heričko, M., Polančič, G.: An empirical investigation of intuitive understandability of process diagrams. Comput. Stand. Interfaces **48**, 90–111 (2016)
15. Van der Aalst, W.M.P., Bichler, M., Heinzl, A.: Robotic process automation. Bus. Inf. Syst. Eng. **60**, 269–272 (2018)
16. Aguirre, S., Rodriguez, A.: Automation of a business process using Robotic Process Automation (RPA): a case study. In: Figueroa-García, J.C., López-Santana, E.R., Villa-Ramírez, J.L., Ferro-Escobar, R. (eds.) Communications in Computer and Information Science, pp. 65–71. Springer, Cham (2017)
17. Kroll, C., Bujak, A., Darius, V., et al.: Robotic process automation - robots conquer business processes in back offices. Capgemini Consult., 1–48 (2016). https://www.capgemini.com/consulting-de/wp-content/uploads/sites/32/2017/08/robotic-process-automation-study.pdf
18. Allweyer, T.: Robotic process automation – Neue Perspektiven für die Prozessautomatisierung (2016). https://www.kurze-prozesse.de/blog/wp-content/uploads/2016/11/Neue-Perspektiven-durch-Robotic-Process-Automation.pdf. Accessed 10 Jan 2019
19. Anagnoste, S.: Robotic Automation Process - the next major revolution in terms of back office operations improvement. Proc. Int. Conf. Bus. Excell. **11**, 676–686 (2017)
20. Moreno-Montes De Oca, I., Snoeck, M., Reijers, H.A., Rodríguez-Morffi, A.: A systematic literature review of studies on business process modeling quality. Inf. Softw. Technol. **58**, 187–205 (2015)
21. De Weerdt, J., Schupp, A., Vanderloock, A., Baesens, B.: Process Mining for the multi-faceted analysis of business processes - a case study in a financial services organization. Comput. Ind. **64**, 57–67 (2013)
22. Czarnecki, C., Auth, G.: Prozessdigitalisierung durch robotic process automation. In: Barton, T., Müller, C., Seel, C. (eds.) Digitalisierung in Unternehmen. AW, pp. 113–131. Springer, Wiesbaden (2018)
23. Manutiu, S.: Digitalisierung im controlling - mehrwert durch robotic process automation. Controlling **30**, 4–10 (2018)
24. Tornbohm, C., Dunie, R.: Market guide for robotic process automation software - Gartner Inc., December 2017. http://images.abbyy.com/India/market_guide_for_robotic_pro_319864(002).pdf. Accessed 10 Jan 2019
25. Forrester Research: Building a Center of Expertise to Support Robotic Automation: Preparing for the Life Cycle of Business Change (2014). http://neoops.com/wp-content/uploads/2014/03/Forrester-RA-COE.pdf. Accessed 12 Jan 2019
26. Willcocks, L., Lacity, M., Craig, A.: The IT Function and Robotic Process Automation: Research on Business Services Automation. Outsourcing Working Research Paper Series, pp. 1–39 (2015)

27. Škrinjar, R., Trkman, P.: Increasing process orientation with business process management: critical practices'. Int. J. Inf. Manag. **33**, 48–60 (2013)
28. Scheer, A.-W.: Robotic Process Automation (RPA) revolution der unternehmenssoftware. IM+io Mag. Innov. Organ Manag. **32**, 30–41 (2017)
29. Hermann, K., Stoi, R., Wolf, B.: Robotic process automation im finance & controlling der Mann+Hummel Gruppe. Control Z. Erfolgsorientierte Unternehm. Sführung **3**, 28–34 (2018)
30. Sachdeva, P.: The Combined Power of RPA and BPM (2018). https://princetonblue.com/combined-power-rpa-bpm/. Accessed 10 Jan 2019
31. UiPath: UiPath Studio, software robot for business automation (2019). https://www.uipath.com/product/studio. Accessed 11 Feb 2019

Sustainable Supply Chain Management

Current Trends in B2C E-Commerce Logistics – A Content Analysis

Christian Straubert(✉), Björn Asdecker, and Immanuel Zitzmann

Chair of Operations Management and Logistics,
University of Bamberg, Bamberg, Germany
Christian.Straubert@Uni-Bamberg.de

Abstract. In the B2C e-commerce market, logistics plays a crucial role, not only as a cost factor, but also as a major success factor for the firms. It is therefore important for e-tailers to know which trends in B2C e-commerce logistics are considered to be the way forward. This also has implications for the directions of future research.

We present a systematic content analysis of 87 non-scientific, practice-oriented articles published on the internet. This results in a comprehensive overview of the trends currently being discussed in B2C e-commerce logistics. An additional correlation/association analysis reveals important relationships between the trends.

A total of 36 trends were identified. Overall, the trend towards faster deliveries was most frequently mentioned in the articles. Followed by more transparency across the processes (track-and-trace), logistics cooperation/outsourcing and more (smaller) urban warehouses.

To the best of our knowledge, this is the first scientific paper that systematically examines current trends in B2C e-commerce logistics practice. The results of this article can serve as an impetus for a variety of research questions, some of which we will touch upon in the course of this article.

Keywords: e-commerce logistics · Future · Trend · Content analysis

1 Introduction

The B2C e-commerce sector has been growing significantly over the recent years and in the search for ways to differentiate their service, many e-tailers identified logistics as a central part of their value proposition to the customer. The delivery of the ordered goods is becoming faster (a recent trend being same-day delivery) and more delivery options (parcel lockers, time-window delivery) are available.

However, this trend comes at a cost. For the year 2016 (the last time the industry leader Amazon.com, Inc. published its net shipping cost) AMAZON.COM (2017) reported that its shipping costs (∼12% of all operating expenses) were roughly 1.78 times higher than its shipping revenues (delivery fees). In addition, there are fulfillment costs (storage, packaging, etc.) which make the whole logistics operations even more expensive.

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 123–140, 2019.
https://doi.org/10.1007/978-3-030-29821-0_9

In every recent report, Amazon.com, Inc. explicitly discusses logistics as a major risk but also as a major success factor for the company. China's largest e-commerce firm JD.com, Inc. names similar risks and places an additional emphasis on the strategic role of logistics for its business model. Otto GmbH & Co KG, Germany's biggest e-tailer, is no different. This is illustrated by the following selected quotes:

"We expect our cost of shipping to continue to increase to the extent our customers accept and use our shipping offers at an increasing rate, we reduce shipping rates, we use more expensive shipping methods, and we offer additional services."

"If we do not adequately predict customer demand or otherwise optimize and operate our fulfillment network and data centers successfully, it could result in excess or insufficient fulfillment [...] In addition, a failure to optimize inventory in our fulfillment network will increase our net shipping cost by requiring long-zone or partial shipments."

(AMAZON.COM 2018, p. 27 and p. 37)

"We have adopted shipping policies that do not necessarily pass the full cost of shipping on to our customers. We also have adopted customer-friendly return and exchange policies that make it convenient and easy for customers to change their minds after completing purchases. [...] These policies improve customers' shopping experience and promote customer loyalty, which in turn help us acquire and retain customers."

"We deliver a majority of the orders directly to customers ourselves, and therefore our customers interact with delivery personnel more often than with any other representatives of our company. [...] We believe that our professionally trained delivery personnel are important in helping us to shape customer experience and distinguish ourselves from our competitors. [...]"

(JD.com 2018, p. 21 and p. 63)

"The logistics area occupies a key position within the Otto Group. Highly advanced processes and systems are employed [...]. However, modified business models for intra-Group customers are leading to new logistical challenges. [...] For this reason, logistical and IT changes must be made to existing systems [...]."

"WE MUST DELIVER! The customers increasingly want to determine exactly how, when and where they receive their parcels. Logistics providers additionally have to deal with debates around congestion and poor air quality in cities; zero-emission transport and even car-free cities are no longer theoretical concepts."

(Otto GmbH & Co KG 2018, p. 105 and p. 40)

Based on these statements it becomes evident that logistics plays a major role for e-commerce firms. The logistics services offered by an e-tailer are essential to their business model and long-term survival. Logistics is treated as a strategic value proposition, which is used to create a competitive advantage. At the same time, logistics is a major operational cost driver. E-tailers have to decide which logistics strategy they want to pursue. Which delivery options should they offer at which prices? Which trends can and which trends cannot be ignored? Currently, only very limited scientific literature exists that covers this topic. Indeed, to the best of our knowledge, a comprehensive discussion does not exist. This paper is intended to close one part of this research gap by using a content analysis of various sources (articles, white papers) published on the internet. The goal is to have a comprehensive overview of what trends are currently being discussed. This also has implications for the scientific community in that it indicates which trends should be researched, as there is a need in practice. Thus, our research question is:

– What are the current trends in B2C e-commerce logistics?

2 Existing Literature

Discussions about the future of logistics have a tradition. There are some contributions concerning the general future of logistics (e.g. Ballou (2007)). Other papers focus on more specific topics such as Industry 4.0 (e.g. Hofmann and Rüsch (2017)), Reverse Logistics (e.g. Govindan et al. (2015)) or City Logistics (e.g. Taniguchi et al. (2014)). However, many papers about the general future of logistics are rather subjective and only very few papers explicitly discuss the future of B2C e-commerce logistics.

Gracht and Darkow (2010) used a comprehensive Delphi study in order to identify probable trends for the logistics services industry. Their research questions are not e-commerce specific and rather general: "How will the macro-environment (political/legal, economic, socio-cultural, and technological structure) of the logistics services industry change by 2025?" and "How will the micro-environment (industrial structure) of the logistics services industry change by 2025?" (Gracht and Darkow 2010, p. 46). Two of the identified scenarios may be of interest in the context of B2C e-commerce: "Customer demands for convenience, simplicity, promptness, and flexibility have turned logistics into a decisive success factor for customer retention" and "Alternative distribution networks have been established in the CEP-market (courier, express, parcel). Petrol stations, kiosks, and local public transport are increasingly used for pickup and delivery of parcels." (Gracht and Darkow 2010, p. 54).

Bask et al. (2012) performed a literature review of e-commerce logistics research and discuss topics for future research. Their focus is B2C, B2B and C2C e-commerce. Some findings are of relevance to the paper at hand. For example, delivery times play a decisive role in the choice between the online channel and brick-and-mortar retail (Bask et al. 2012, p. 13). However, Bask et al. (2012) found that only a few scientific articles discuss the matter hand (i.e. B2C e-commerce logistics): "Based on the review, however, it seems that the logistics solutions have not been as extensively analyzed as could be expected, given that physical delivery is a key issue in successful e-commerce. In many cases, e-commerce was approached from the information systems or marketing points of view, and thus not many specific logistics solutions for e-commerce deliveries were proposed or analyzed." (Bask et al. 2012, p. 10).

Based on the literature search we conclude that there is a research gap with regard to B2C e-commerce logistics. Papers such as Ballou (2007) are rather abstract, while the paper at hand is much more specific but comprehensive at the same time. Gracht and Darkow (2010) is a methodically sound Delphi study but not about e-commerce. Bask et al. (2012) has a rather similar objective compared to the paper at hand but is based on scientific literature. Our paper is predominantly practice-oriented. Bask et al. (2012) emphasize that the scientific literature is lacking in the area of e-commerce logistics, indicating that a more practice-oriented approach might be better to identify the current trends in B2C e-commerce logistics.

3 Methodology

In order to identify the trends, we used a 'Problem-Driven Content Analysis' (Krippendorff 2013, pp. 357–370). Besides the fact that there is very little scientific literature on this topic, we decided to use mostly non-scientific online articles as source material because the topic is practice-oriented and the internet is known as a fast, responsive publication channel.

Although the content analysis method has rarely been used in logistics research, it fits the purpose of this study. Mir et al. (2018), p. 166, who did a review of content analysis in logistics research, explicitly state: *"quantitative content analysis methods may be utilized to examine the latest trends in last mile delivery such as Amazon's intentions to use drones for order fulfillment."*

Krippendorff (2013), p. 358, named nine general steps to consider for a methodologically sound, problem-driven content analysis:

1. Formulating research questions
2. Ascertaining stable correlations
3. Locating relevant texts
4. Defining and identifying relevant units in texts
5. Sampling these units of texts
6. Developing coding categories and recording instructions
7. Selecting an analytical procedure
8. Adopting standards
9. Allocating resources

Due to the nature of our analysis some steps are of lesser/minor relevance. We do not need to 'ascertain stable correlations' because we are mainly interested in what is said literally. Furthermore (similar to Spens and Kovács (2006)), we can mostly omit the steps 'Defining and identifying relevant units in texts' and 'Sampling these units of texts'. Most of the publications found on the Internet are short articles that can be read in their entirety. We split the articles so that each paragraph, heading and bullet point is a text unit. During the recording process all these text units were coded one by one. After the recording we mainly used tabulations and statistical tests on correlation/differences for the 'analytical procedure'. The remaining five steps suggested by Krippendorff (2013) are described in the following:

Formulating research questions – Our Content Analysis follows the broad research question: 'Which current B2C e-commerce logistics trends are mentioned on the internet?'

Locating relevant texts – We chose the three search engines 'www.google.com', 'www.bing.com' and 'www.duckduckgo.com' (a meta search engine that does not personalize the search results) to find the relevant texts. In order to find the best search queries we conducted a pre-test with several different search queries such as "future e commerce supply chain", "developments B2C ecommerce logistics" or "trends e-tail logistics". For the final search, we used the following search queries because they yielded by far the best search results:

"future e-commerce logistics"
"trends e-commerce logistics"

For each search engine and query, we checked every search result until we reached a point at which 20 consecutive search results were irrelevant. We performed the search between the 2018-08-29 and the 2018-09-04. We recorded only texts that are written from the viewpoint of the most mature e-commerce markets in the world (Europe, China, North America, ...). Furthermore, we only considered publications dating from 2014 or later in order to focus on sources that are more likely to describe current trends in B2C e-commerce logistics and because the search restricted to those years yielded enough results. It was therefore not necessary to search further in the past. After the final recording, we calculated the *Pearson correlation coefficients* for the trend-year combinations and found no clear correlations between the year of the source and the trends. This suggests that the articles from 2014 are not outdated.

During the screening process of the query results we paid attention if a text prominently featured an aspect of the two themes 'future/trends of e-commerce' and/or 'future/trends of e-commerce logistics' either in the title, abstract or introductory paragraph. If a text was deemed relevant for further analysis, it was read (in most cases in its entirety). If the text indeed contained content about the current trends in B2C e-commerce logistics, it was logged in a spreadsheet that we used during the screening process. We used the following definition of B2C e-commerce logistics during this step and the actual recording process: The physical scope of B2C e-commerce logistics is every touch-point the finished product has after it was ordered by the customer. In addition to the physical interactions (warehouse, delivery, ...), the information flows associated with these interactions and the configuration of the network/logistics market are also included. We considered a topic to be a trend if the author explicitly mentions that it will become more important/increase in the future or if the topic/technology/method did not exist or was less prevalent in the past.

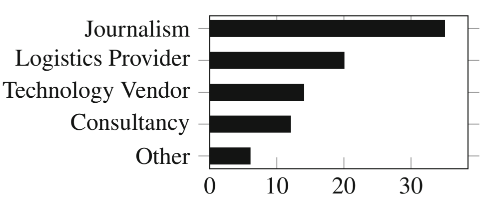


Fig. 1. # of texts by type of source

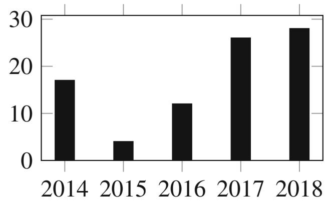


Fig. 2. # of texts by year

In total 87 texts were recorded. Figure 1 shows the distribution of the texts by type of source (e.g. Journalism) and Fig. 2 shows the number of texts aggregated by year. It is noteworthy that relatively few relevant articles were found for the year 2015 (especially compared to the year 2014). If one assumes that the number of hits decreases from year to year, the further one goes into the past (because the search engines find older articles to be less relevant), it is surprising that so many sources were found for the year 2014 compared to so few for 2015. However, this outlier for 2014 can be explained by a large-scale study by DHL which was published in 2014. In this study, four different scenarios are presented, which are considered separately in the content analysis. In addition, the study contains eight expert interviews, which are also separately included in the content analysis.

Developing coding categories and recording instructions – The coding process itself was challenging. It is an explorative content analysis. This means that the codes/trends are not known at the beginning of the analysis. Also many trends/concepts are named in multiple sources but with slightly different terminology (e.g. 'autonomous shipping vehicles' vs. 'self-driving shipping vehicles'). Therefore, in a first phase, two coders A and B separately looked at all text units and tried to identify all mentioned trends. Then the two coders harmonized the terminology and created a codebook. Particular care was taken to ensure that the codes are exhaustive and mutual exclusive. The codebook (including coding instructions) was iteratively refined based on a test data sample until a satisfactory inter-coder reliability was achieved (Krippendorff 2013, p. 130).

Adopting standards and Allocating resources – In order to minimize a bias that could occur if all coders were very familiar with the codebook (Krippendorff 2013, p. 131) coder B who took part in the first phase was swapped with a third coder C who was not involved in the process of creating the codebook. Coder C was trained on a small test data sample (5 articles). In the second phase coder A and coder C coded all text units based on the codebook created in the first phase. This coding process was difficult because many different trends were identified in the first phase and often several different trends were mentioned in one text unit. In order to make the large number of codes easier to handle, a hierarchy system was used (Krippendorff 2013, p. 135).

Table 1. Inter-coder reliability scores

Score	Value	LowerCB 95%	Upper confidence bound 95%
Agreement Rate	0.754		
Cohen's Kappa	0.672	0.653	0.691
Scott's Pi	0.672	0.652	0.691
Krippendorff's Alpha	0.672		

Table 1 contains the 'Agreement Rate' (Matching Recordings/All Recordings) and the inter-coder reliability scores 'Scott's Pi', 'Cohen's Kappa' and 'Krippendorff's Alpha'. All values fall within ranges that are regarded as acceptable or good (Neuendorf 2002, p. 143). Whether a certain amount of inter-coder unreliability can be accepted is always dependent on the matter at hand and the research goals. Since our main research goal is to identify the current trends in B2C e-commerce logistics we can accept a moderate level of unreliability because the results are not invalidated if, for example, the actual count of a trend is 60 times versus a recorded count of 70 times. All disagreements were discussed between coder A and coder C until a consensus decision was reached (similar to Mir et al. (2018) and Seuring (2008)).

It is always difficult to make predictions about the future. This is no different in our case either. Nevertheless, the case can be made, that if different people agree on predictions, it is more likely that these predictions come true (Rowe and Wright 2001, p. 128). Furthermore, it stands to be reasoned that if the people who made the predictions can influence the subject matter, that the predictions become a self-fulfilling prophecy. In our case, the people who made the predictions are journalists, logistics/technology providers and consultants, all of whom have the ability to influence the matter at hand. The journalists influence the public. The logistics and technology providers influence the e-commerce firms through their products. And the consultants influence the e-commerce firms because they ask the consultants for advice. The nature of our analysis is rather straightforward and therefore less vulnerable in general. The use of three different search engines makes it probable that most of the relevant texts have been identified. Therefore, we are confident that our analysis, which we present in the following, is valid.

4 Results of the Content Analysis

Table 2 shows how often the trends were recorded. For better readability some trends that were mentioned only a few times are not included in the table. We grouped the trends under headings like 'Analytics/Software' and 'Types of Delivery' for a better understanding of common themes. Please note that one article can contain multiple trends.

We also performed an additional correlation (or rather association) analysis, the results of which can be found in Table 3. We first counted how often a pair of

Table 2. Number of times a trend was named (# of articles)

Trend	# of articles	% of articles
Faster Delivery	72	83%
Generally Faster Delivery	59	68%
Same-Day Delivery	34	39%
One- /Two-Hour Delivery	15	17%
Types Of Delivery	52	60%
Click-and-Collect/Parcel Locker	28	32%
Drone Delivery	24	28%
General More Delivery Options/Variety	18	21%
Time Window Delivery	9	10%
More Delivery Hours	7	8%
Anywhere/Any-Place Delivery	5	6%
Delivery Characteristics	50	57%
Green Logistics	17	20%
Free/Low Price Delivery	16	18%
Easier/Free Returns, More Returns	16	18%
Value Added Services/Quality in Delivery	13	15%
Difficult Deliveries (Food Delivery, ...)	11	13%
Bundling, City Logistics Concepts	9	10%
Delivery Fee Subscriptions	5	6%
Periodic Deliveries (e.g. 10th of each month)	3	3%
Analytics/Software	49	56%
Data/Operations/Process Transparency	36	41%
More/Better Software/Analytics	25	29%
Big Data/Business Intelligence	19	22%
Machine Learning/Artificial Intelligence	18	21%
Predictive Shipping/Logistics	10	11%
Automation, Technology in Warehouse/Delivery	41	47%
Warehouse/Fulfillment Automation	20	23%
Better Technology in Warehouse/Fulfillment	17	20%
Automated Delivery Vehicles	17	20%
General Automation	14	16%
Better Technology in Delivery	10	11%
Logistics Market/Infrastructure	66	76%
Logistics Outsourcing/Logistics Cooperation	32	37%
(Smaller) Regional/Urban Warehouses	30	34%
Multi-Channel/Omni-Channel	28	32%
Logistics Insourcing	24	28%
Logistics Marketplaces, Crowd Logistics	20	23%
International/Cross-Border Logistics	16	18%
Bigger Warehouse/Fulfillment Infrastructure	7	8%
Mobile Warehouses (Store on Wheels, ...)	5	6%

Table 3. Significant associations between trends – often occurring combinations in the articles

Trend 1	Trend 2	# of articles	% of articles	Cramer's V
Faster Delivery	(Smaller) Regional/Urban Warehouses	29	33%	0.257*
Data/Operations/Process Transparency	Logistics Outsourcing/Logistics Cooperation	20	23%	0.322**
Data/Operations/Process Transparency	More/Better Software/Analytics	18	21%	0.391***
(Smaller) Regional/Urban Warehouses	Logistics Outsourcing/Logistics Cooperation	17	20%	0.295*
(Smaller) Regional/Urban Warehouses	Click-and-Collect/Parcel Locker	16	18%	0.325**
Multi-Channel/Omni-Channel	Logistics Outsourcing/Logistics Cooperation	16	18%	0.287*
Data/Operations/Process Transparency	Logistics Marketplaces, Crowd Logistics	16	18%	0.426***
(Smaller) Regional/Urban Warehouses	Multi-Channel/Omni-Channel	15	17%	0.272*
Data/Operations/Process Transparency	Warehouse/Fulfillment Automation	14	16%	0.314**
(Smaller) Regional/Urban Warehouses	More/Better Software/Analytics	14	16%	0.284*
Data/Operations/Process Transparency	Better Technology in Warehouse/Fulfillment	14	16%	0.407***
Multi-Channel/Omni-Channel	Click-and-Collect/Parcel Locker	14	16%	0.259*

*/**/*** Significant on 95%, 99% or 99.9% level respectively

trends was found within the articles. Reading example: The first row in Table 3 means that 29 articles contained the trend '*Faster Delivery*' as well as the trend '*(Smaller) Regional/Urban Warehouses*'. Then we calculated Cramér's V, a measure of association between the two trends (nominal variables; trend mentioned: yes/no), for each trend pair. We only report trend pairs that are significant on at least a 95% confidence level. We restricted our analysis to trend pairs that were found in at least 13 articles (∼15% of all articles) in order to ensure the generalizability of the results. All reported trend pairs have a moderate (0.20 to <0.40) or relatively strong association (0.40 to <0.60) (Rea and Parker 2014, p. 219).

We examined each association by studying the underlying set of articles. In the following discussion, the results of the statistical analysis will be underpinned and substantiated by selected quotes from the recorded texts.

Overall, the trend towards faster deliveries was most frequently mentioned. *Same-Day Delivery* or even *One-/Two Hour Delivery* was often explicitly mentioned as the goal while *Slower Delivery* (which is not listed in the table) was only mentioned three times. This addresses one of the last advantages of brick-and-mortar retail: instant gratification. The customers want fast delivery (at least as an option) and while it makes the delivery more expensive it also provides the opportunity for the e-tailers to shift revenue from traditional brick-and-mortar retail towards e-commerce. This aspect is also one of the findings of the literature review performed by Bask et al. (2012) (see Sect. 2), which indicates that this is still one of the most important trends.

However, the last mile was and is a problem that will only become more acute due to increased demands in terms of speed and flexibility. Some articles predict that the very costly *Time-Window Delivery* option will become more popular. Other articles say that there will be more hours of operations (e.g. delivery on Sundays or in the early morning, late evening) and that it will be possible to receive the parcel anywhere. This increases the costs for the last mile, but *Drone Delivery*, *Click-and-Collect/Pick-Up and Parcel Lockers* might be viable solutions to this problem. Drones are mostly automated, not bound to the (possibly congested) road network and can travel in linear distance. This greatly reduces personnel costs and can shorten delivery times. *Click-and-Collect* or *Parcel Lockers* basically outsource the burden of the last mile to the customer. But at the same time they provide added value to those customers who want to decide when to receive their parcel. Not listed in the table are *In-Car Delivery* and *In-Home Delivery* which were named only in two or one article/s respectively.

Damian Harrington (2015), Colliers International

"As the essential element in improving urban logistics is to limit deliveries to the shortest route, e-commerce retailers have started to include smaller urban warehouses to shorten delivery routes and provide quick delivery services to online customers."

If the future indeed lies in very fast delivery it is necessary to build more (smaller) warehouses that are close to the urban demand. This is a clearly named trend that was mentioned in many articles and is emphasized by our statistic: the trend towards more *(smaller) regional/urban warehouses* was named in 34% of all articles compared to more *big warehouses* which was predicted in only 8% of all articles. In addition, the calculation of Cramér's V reveals that the association between the trend '*Faster Delivery*' and '*(Smaller) Regional/Urban Warehouses*' is indeed statistically significant. The next most frequent trend pair involving '*(Smaller) Regional/Urban Warehouses*' is with '*Logistics Outsourcing/Logistics Cooperation*' highlighting that many e-commerce firms, especially the smaller ones, will struggle to establish their own network of regional/urban warehouses. This is exemplified by the following quote:

Michael Lierow et al. (2014), Oliver Wyman

"When it comes to innovative logistics solutions, pure-play SMEs may find introducing a physical means of differentiation such as same-day delivery tough to implement, as most have only one or a few centralized warehouses and thus cannot offer broad same-day delivery. Once same-day takes off (and customers demand it), solutions could include developing more decentralized distribution centers or teaming up with other niche SMEs to set up local same-day warehouses for critical stock. Enabling IT won't be trivial [...]"

The above text excerpt also notes that there are new IT challenges associated with small urban warehouses. This association between the trends '*(Smaller) Regional/Urban Warehouses*' and '*More/Better Software/Analytics*' is also statistically significant in our sample.

We see great research potential in this area. Many small e-commerce firms struggle to compete with the big industry leaders such as Amazon.com. Some pundits even consider the market an oligopoly. Trends like faster delivery will make the competitive edge of the industry leaders even more pronounced. Small businesses will probably have no other choice than to cooperate with one and another, either through an intermediate 3PL provider or directly. Urban Warehouses have to be shared facilities in order to achieve the necessary throughput. But why stop there? Many small firms will also struggle to even have stock at every urban warehouse. They would have to increase their number of stock keeping units and suffer an increase in tied-up capital. It will therefore probably be necessary for small firms to cooperate with inventory sharing whenever possible and sensible. The necessary coordination is indeed not trivial and we are planning further research on this topic.

Brian Straight (2018) for FreightWaves

"Approximately three-quarters of retailers (76%) currently use store inventory to fill online orders with 86% planning to institute buy online/pick up in store programs next year. Some retailers are retrofitting stores to double as online fulfillment centers and 70% of surveyed executives believe this trend will grow."

Some authors argue that the demand for big warehouses will increase or at least remain steady for the next years due to many traditional brick-and-mortar retailers who still do not have a comprehensive e-commerce strategy. But the shift towards a *multi-channel* model is inevitable for many of these retailers because they would otherwise have trouble to survive the increased competition. However, the brick-and-mortar retailers are in a good position to make the shift because they already have a lot of *urban warehouses*: their stores. Several articles argue that these stores will become multi-use facilities: A traditional place to shop, a click-and-collect point and a mini-fulfillment center for urban demand. This was also a result of the Delphi study performed by Gracht and Darkow (2010) (see Sect. 2).

But the shift from traditional retail towards a multi-channel model is not easy, which is why a lot of retailers have to rely on 3PL service providers, at least in the beginning.

Wenda Ma (2017), HKTDC Research

"The retail industry is undergoing a structural change amid the proliferation of e-commerce – increasingly applying an online-to-offline model to their businesses. The new mandate for omni-retailers includes customer-centricity, digital fluency, and complete agility. Rather than developing their own logistics networks and delivery systems, relying on a 3PL provider's extensive distribution and delivery networks might help retailers that are currently behind the curve to rapidly establish an online presence."

Big industry leaders like Amazon.com on the other hand are increasingly in-sourcing their logistics operations. This gives them more control over the processes and provides them with an additional opportunity to differentiate themselves from the competition. Of course faster delivery is not the only viable strategy. Many authors also argue that *Cross-Border Logistics* will increase. This is a very natural trend (globalization) and the technical barriers (website) are low. Additionally, this strategy increases the economies of scale which make it possible to offer the products at a lower price (if the increased shipping costs do not overcompensate). Already today business models exist (e.g. AliExpress.com), where the products are directly shipped from China to customers in Europe or the USA.

Another widely discussed topic is the increased use of advanced *Software and Analytics*. In this theme the most frequent topic which was covered in 41% of all articles is *Process Transparency*, either with regard to the parcel (Track-and-Trace) or with regard to the internal processes (e.g. stock levels). Many articles featured popular topics such as *Big Data* and *Artificial Intelligence*. These technologies are trending in many areas and it is very likely that they will also play an important role in B2C e-commerce logistics. One particularly difficult application of these technologies is *Predictive Shipping/Logistics* which describes a process where selected products are picked from the warehouse shelf without a corresponding order. In some cases this product is also loaded onto a truck anticipating that some customer in the respective delivery area will order the product in the next hours. Not listed in the table is the *Blockchain* technology which was only named in two articles. All in all, *Software and Analytics* are classic enablers that help to implement other trends. Be it the coordination of different logistics service providers (freight brokerage), the coordination of many different independent 'Uber-like' courier drivers (who enable faster delivery to customers) via *Logistics Marketplaces*, or the increased use of (better) *Automation and Technology* in warehouse and delivery processes.

Charlie Hitt cited by John Schulz (2018) for Logistics Management

"'The big differentiator is technology' Hitt explains. 'Everybody at XPO looks at technology as a leading driver as to what we do with customers. Technology manages workflows, directs the customer experience and provides visibility to all parties, customers and retailers.'"

Better Technology in Warehouses and Delivery like Radio-Frequency Identification and mobile devices connected to the internet/intranet is necessary in order to make the processes transparent. Other technology makes the pick and pack process faster (e.g. pick-by-light systems in warehouses) or help to increase the quality of the delivery process (e.g. sensors in the packages). The advantage of *Warehouse Automation* and *Automated Delivery Vehicles* is evident in that the technologies help to reduce the payroll costs and make the process less vulnerable to fluctuations in capacity and quality.

However, a more automated and therefore anonymous delivery process also has disadvantages. As JD.com puts it (see introduction above): *"[...] our customers interact with delivery personnel more often than with any other representatives of our company. [...] We believe that our professionally trained delivery personnel are important in helping us to shape customer experience and distinguish ourselves from our competitors."* Several articles talked about the trend towards increased *quality in delivery* (e.g. White Glove Delivery) to increase the value of the service and differentiate it from other more standard delivery experiences. Another possibility for differentiation are the return policies of the e-commerce firms. Several articles predict that the returns management will become even more important and an easy returns process can help to distinguish a firms' service offering from the competitors. One could argue that *Green Logistics* concepts also have the potential to make a delivery service special, but this is probably not the case in the future. While some articles see it as a way of differentiation, most articles simply state that this a general trend and many large e-tailers and logistics companies have programs to reduce the CO_2 output and air pollution. It will basically become obligatory and the only difference is the scope and number of measures taken. Another basic trend is that the *customer does not have to pay a delivery fee* or only has to pay a subsidized fee that is lower than the actual costs. Some firms use *Delivery Fee Subscriptions* (e.g. Amazon Prime) in order to provide the customer with unlimited delivery for a fixed fee. Delivery fee subscriptions also play an important role in the *difficult grocery e-commerce*. After the failure to establish grocery e-commerce during the dot-com bubble (e.g. 'Webvan') some authors think that the time (demand and technology) has now come for e-commerce to establish itself in the food sector. Other authors remain skeptical due to the immense costs associated with the delivery of refrigerated goods. However, it is a natural assumption (although not necessarily correct) that difficult deliveries (also furniture, luxury items, etc.) will increase as these areas still have a comparatively small e-commerce market share and a lot of start-ups are currently emerging who see their chance.

Looking at the big picture, one might ask whether it is possible to predict which trends are more likely to occur than others based on observations other than the mere frequency with which a trend was mentioned. Two possible considerations are:

(1) Some trends are ongoing, i.e. they existed in the past, but the change process is not yet finished (an example for this is the trend towards faster delivery). Other trends are mere predictions of changes that may occur in the future but are currently at an early stage or have not even been implemented outside of field trials. It can be assumed that trends of the first category are more likely to exist in the future than trends of the second category. For trends of the second category, uncertainty is greater due to a lack of actual experience.

(2) The identified trends can be split into actual goals and methods/technologies which help to achieve these goals. For example faster delivery is an actual goal and smaller urban warehouses are a method/technology that enables the e-tailers to provide fast delivery because the inventory is stored close to the customer. Some of these relationships are also emphasized by the presented association analysis. It can be assumed that methods/technologies that help to achieve one or several goals without conflicting with the other goals have an increased chance of being successful. Lower-level goals such as faster delivery will be more successful if they support higher-level goals such as financial profit or customer satisfaction. The separation between goals and methods/technologies is gradual, since every method can be made a goal and lower-level goals can also be interpreted as methods to achieve high-level goals.

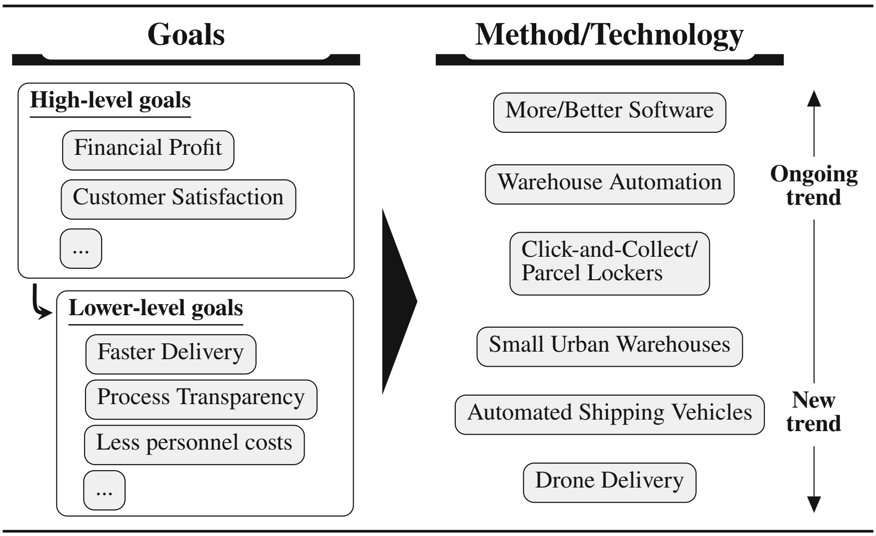


Fig. 3. Goals versus methods/technologies – A framework for the classification of trends

Figure 3 illustrates these viewpoints with some of the most frequently mentioned trends. These two considerations may also be of value for the assessment of trends that do not appear in this study but may arise in the future. A theoretical framework that allows the assessment of (new) trends in B2C e-commerce logistics in general is a desirable goal for future research efforts.

5 Conclusion and Outlook

The paper at hand incorporates both a scientific and a practical perspective. Our content analysis helps to put the ubiquitous discussions about the future of B2C e-commerce logistics on a sound footing that can be used as a basis for further research, or, if we look further into the future, as a snapshot in time that can serve as a basis for comparison. The results provide strong indications as to which trends are promising research areas and which are not. For example, almost no article mentions *In-Car Delivery*, *In-Home Delivery* and *Blockchain* as promising trends, suggesting that these concepts are a dead end.

A central finding of our study is that often faster delivery is considered the central goal for the future. Many of the other trends mentioned by the articles can be connected to this. Other central goals are transparency about the order/delivery status, more quality during the process, environmentally friendly operations and more service/convenience for the customer. It can be assumed that if a method/technology supports several of these goals without conflicting with the other goals or high-level objectives such as financial profit, it has an increased chance of being successful. Methods and technologies often mentioned are click-and-collect and parcel lockers, drone delivery, more and better software, warehouse automation, logistics outsourcing/insourcing (depending on the situation and also in combination with logistics/driver marketplaces), smaller urban warehouses and a multi-channel business model. Our association analysis emphasizes some of these relationships. For example, the most often occurring statistically significant trend pair is faster delivery together with smaller urban warehouses. If one wants to achieve short delivery times, it is physically necessary to keep the inventory close to the customer. Another viewpoint when evaluating the likelihood that a trend will occur is the age of said trend. If a trend exists for some time, but still has room for more change, it is a reasonable assumption that this trend will continue for the foreseeable future. Faster delivery is one example of an ongoing trend.

Basically, it is quite natural that companies try to reduce their costs and improve their service through, for example, more automation or better software, and thereby create a competitive advantage over their competitors. However, the methods employed can differ from firm to firm and our content analysis may shed some light onto which methods are currently considered to be the way forward.

Identifying the current trends in B2C e-commerce logistics and their relationships, however, is only the first step towards fully exploring the topic. Research is needed from multiple angles. Many of the trends have implications regarding the planning and execution of last mile delivery. Disrupting technologies such as drone delivery or automated delivery vehicles require other planning priorities than the traditional *Vehicle Routing Problem*. The industry leaders increasingly use advanced methods such as machine learning to better predict customer demand and the IT-systems of all firms are more and more interconnected. This opens up new possibilities like inventory sharing between multiple companies, which is often necessary to provide very fast delivery without exploding costs, but also has more generally the potential to increase welfare by making the economy more efficient.

The e-commerce firms have to decide how they use the information provided in this paper. As we said in the introduction of the paper: E-tailers have to decide which logistics strategy they want to pursue. Which delivery options should they offer at which prices? Which trends can be and which trends cannot be ignored based on their logistics strategy? Perhaps one answer to this lies in the classical competitive strategy theories by Porter (Porter 1998). Or it is necessary to develop a B2C e-commerce specific framework (e.g. Ghezzi et al. (2012)). In this paper we present a framework that categorizes the trends in *goals* and *methods/technologies*. This could help to better understand the relationships between the trends and be useful when drafting a logistics strategy. In any case, logistics is a central component of the competitive strategy of e-commerce firms and every firm needs a plan how to position itself in the market.

However, the paper has some limitations. Generally, predictions about the future are not deterministic but rather an indication. Moreover, is not entirely known how knowledgeable the authors of the sources used in the content analysis really are. Also the inter-coder reliability was not perfect. The paper presents a good starting point for complementary further research which could use a questionnaire or a Delphi study approach. In the course of this it might also be useful to consider different countries/regions or products groups (market characteristics) separately, as it can be assumed that customer preferences and B2C e-commerce logistics systems may be very different depending on the circumstances. Every market has different characteristics and for a more in-depth look it would be useful to understand these characteristics and correlate them with the different trends. A desirable goal of such further research could certainly be to understand more thoroughly why the trends mentioned in this paper are taking place.

We hope that this paper is useful in guiding researchers and practitioners alike. The B2C e-commerce market is huge and still growing. The underlying logistics and service processes directly impact our quality of life. Thus, the importance of research in this field and its application cannot be underestimated.

References

AMAZON.COM, INC.: Form 10-K - Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934 for the fiscal year ended 31 December 2017 (2017)

AMAZON.COM, INC.: Form 10-Q - Quarterly Report Pursuant to Section 13 or 15(d) of the securities exchange act of 1934 for the quarterly period ended 30 June 2018 (2018)

Ballou, R.: The evolution and future of logistics and supply chain management. Eur. Bus. Rev. **19**(4), 332–348 (2007). https://doi.org/10.1108/09555340710760152

Bask, A., Lipponen, M., Tinnilä, M.: E-commerce logistics: a literature research review and topics for future research. Int. J. E-Services Mob. Appl. **4**(3), 1–22 (2012). https://doi.org/10.4018/jesma.2012070101

Ghezzi, A., Mangiaracina, R., Perego, A.: Shaping the e-commerce logistics strategy. a decision framework. Int. J. Eng. Bus. Manag. **4**(2), 1–13 (2012). https://doi.org/10.5772/51647

Govindan, K., Soleiman, H., Kannan, D.: Reverse logistics and closed-loop supply chain: a comprehensive review to explore the future. Eur. J. Oper. Res. **240**(3), 603–626 (2015). https://doi.org/10.1016/j.ejor.2014.07.012

Gracht, H., Darkow, I.L.: Scenarios for the logistics services industry: a delphi-based analysis for 2025. Int. J. Prod. Econ. **127**(1), 46–59 (2010). https://doi.org/10.1016/j.ijpe.2010.04.013

Harrington, D.: E-commerce and new delivery paradigms shape future of global logistics (2015). https://www.supplychain247.com/article/ecommerce_and_new_delivery_paradigms_shape_future_of_global_logistics. Accessed 26 Feb 2019

Hofmann, E., Rüsch, M.: Industry 4.0 and the current status as well as future prospects on logistics. Comput. Ind. **89**, 23–34 (2017). https://doi.org/10.1016/j.compind.2017.04.002

JD.com, Inc.: Form 20-F - Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934 for the fiscal year ended 31 December 2018 (2018)

Krippendorff, K.: Content Analysis. An Introduction to its Methodology, 3rd edn. SAGE, Los Angeles (2013)

Lierow, M., Janssen, S., D'Inca, J.: Disruptive logistics - the new frontier for e-commerce (2014). https://www.oliverwyman.com/content/dam/oliver-wyman/global/en/2014/sep/MUN-MKT20101-011_screen12.pdf

Ma, W.: E-commerce: the logistics industry (2017). http://hkmb.hktdc.com/en/1X0A9I7B/hktdc-research/E-commerce-The-Logistics-Industry. Accessed 26 Feb 2019

Mir, S., Lu, S.H., Cantor, D., Hofer, C.: Content analysis in SCM research: past uses and future research opportunities. Int. J. Logist. Manag. **29**(1), 152–190 (2018). https://doi.org/10.1108/IJLM-09-2016-0200

Neuendorf, K.: The Content Analysis Guidebook. SAGE, Thousand Oaks (2002)

Otto GmbH & Co KG: Annual Report 2017/18 (2018)

Porter, M.: Competitive Strategy. Techniques for Analyzing Industries and Competitors. Free Press, New York (1998)

Rea, L., Parker, R.: Designing and Conducting Survey Research. A Comprehensive Guide, 4th edn. Wiley, San Francisco (2014)

Rowe, G., Wright, G.: Expert opinions in forecasting: the role of the delphi technique. In: Armstrong, J.S. (ed.) Principles of Forecasting. Springer, Boston, MA (2001)

Schulz, J.: Battle for the last mile. The e-commerce logistics revolution. Logist. Manag. (2018). https://www.logisticsmgmt.com/article/the_e_commerce_logistics_revolution

Seuring, S.: Assessing the rigor of case study research in supply chain management. Supply Chain Manag. Int. J. **13**(2), 128–137 (2008). https://doi.org/10.1108/13598540810860967

Spens, K., Kovács, G.: A content analysis of research approaches in logistics research. Int. J. Phys. Dist. Log. Manage. **36**(5), 374–390 (2006). https://doi.org/10.1108/09600030610676259

Straight, B.: Survey finds manual processes still dominate e-commerce logistics (2018). https://www.freightwaves.com/news/technology/ecommerce-logistics-remains-a-manual-process Accessed 26 Feb 2019

Taniguchi, E., Thompson, R., Yamadaa, T.: Recent trends and innovations in modelling city logistics. Procedia - Soc. Behav. Sci. **125**, 4–14 (2014). https://doi.org/10.1016/j.sbspro.2014.01.1451

Check for updates

Recent Trends in Last Mile Delivery: Impacts of Fast Fulfillment, Parcel Lockers, Electric or Autonomous Vehicles, and More

Martin Behnke(✉)

Martin Luther University Halle-Wittenberg, 06108 Halle, Germany
martin.behnke@wiwi.uni-halle.de

Abstract. This paper presents a review of recent trends in urban fright transportation, especially the typical last mile for logistic service providers. The high increase in shipped parcels over the last and presumably future years makes it necessary to particularly tackle the special logistical issues in large urbanized cities. Of course, one leading factor of this effect is the massive increase in online orders and the resulting deliveries. The advanced delivery concepts, which will be highlighted, try to overcome typical image in most urbanized cities, like congested roads, dangerous parking of delivery vans, high environmental effects, etc. The aim of this work is to present and analyze some of the recent trends in the last mile delivery to show their strengths and weaknesses. Especially, the impact of fast fulfillment offers, parcel lockers or stations, and advanced delivery vehicles like electric vans and bicycles or autonomous vehicles is in the focus.

Keywords: Last mile delivery · Supply chain coordination · Transport management

1 Introduction

Parcel deliveries, especially in the urban context of the last mile problem, have received an increasing amount of reflection over the last years. In [26], a review of the past, present, and expected future trend in transportation and logistics is given together with arising problems and possibilities like the use of electric vehicles, reduction of traveling vehicles and parking space, and reduction of congestion. A more concrete reflection is given in [24], which shows the opportunities from trends and advances in technology impacting city logistics.

The reason for this growing research interest and need tends from the continuously increasing amount of delivered parcels, combined with rising quality and service requirements like fast deliveries, returns, additional services, etc. Consequently, costs for logistic companies increase. On the other hand, the typical image in most urbanized cities results: vans of different companies (DHL, UPS, TNT, DPD, GLS, FedEx, Hermes, just to mention some of them)

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 141–156, 2019.
https://doi.org/10.1007/978-3-030-29821-0_10

are double parking, or in hazardous positions, on footways or bikeways, congesting the traffic, and the greenhouse gas emissions from the required tours - additionally increased by the high amount of short stop and go rides - affects the environment drastically.

The aim of this work is to analyze how the usage of recent last mile concepts can help reducing the effort for logistic companies, and hence their costs, together with the resulting greenhouse gas emissions. The remainder of this work is organized as follows: Sect. 2 presents recent trends and statistics in the field of parcel deliveries. In Sects. 3 and 4, advanced delivery concepts and their influence to the last mile will be discussed and analyzed. Section 5 sums up the findings and gives ideas for future research opportunities.

2 Recent Trends and Statistics in Parcel Delivery

In the recent years, the amount of delivered parcels was increasing drastically. The federal association of parcel and express logistics in Germany reports an increase of 5% shipped parcels per year since five years, see Fig. 1 and they predict that this trend will maintain for at least five more years. One leading factor for this increasing amount of parcel deliveries is the massive increase in online orders and resulting deliveries. The trade association of Germany reports a yearly revenue increase of 10.4% in e-commerce over the last five years, see Fig. 2.

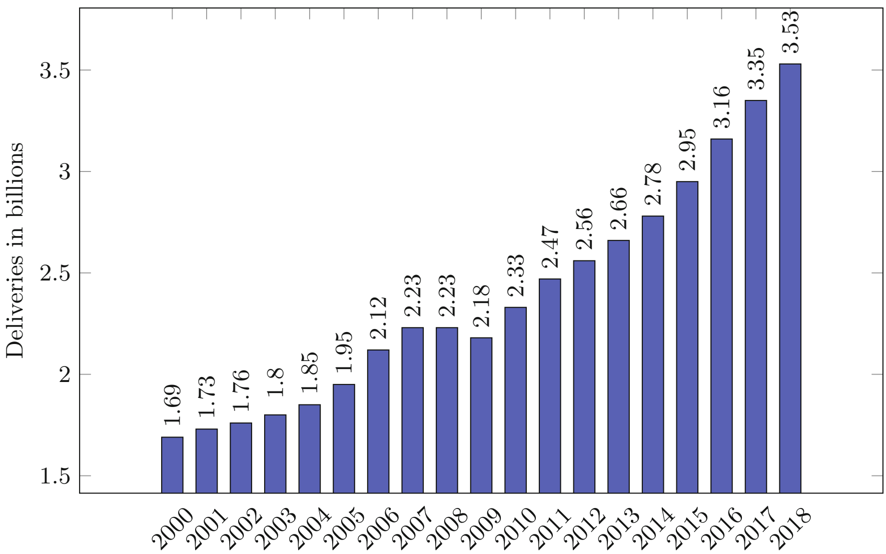


Fig. 1. Courier, express, and parcel deliveries in Germany by year. Own figure with data from [14].

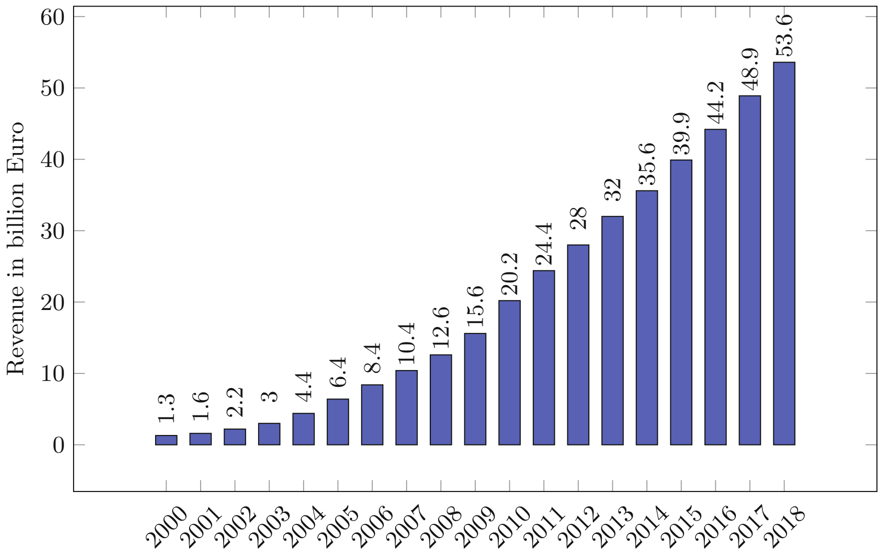


Fig. 2. Revenue from e-commerce in Germany by year. Own figure with data from [6].

Apart from the amount of parcels, also the average delivery time has decreased drastically over years and express deliveries are becoming more and more popular, like next-day or same-day delivery from Amazon [3], DHL [9], FedEx [11], or others, particularly for no additional charge. On the other hand, the additional costs, administrative effort, and the environmental footprint for the retailer and the logistics service provider are not negligible at all. These fast deliveries require a huge logistical effort at all parts of this logistics supply chain. Nevertheless, the amount of qualified areas for these services increases continuously.

Following data from *The Wall Street Journal* and *The Start Business Journal*, the share of returned online purchased products sums up to 20%, see [16], [23], and [25]. This behavior is caused by different reasons: a lot of online shops offer free returns, customers tend to order products in varying style and size and return the unwanted ones, and still damaged parcels or products are relevant.

Another important factor for the overall efficiency of the logistic companies is the failure of delivery trials which plays an important role as it implies additional costs, time, and emissions for the carrier to relocate and handle the parcel additionally. Following different sources (see e.g. [10] and [28]), the first-time failure delivery rates between 2–30% which varies widely among different countries and the carrier's guideline to handle the no-one-at-home scenario. Some logistic companies prepare these situations by considering alternative delivery options like neighbors or specified drop down locations nearby, which of course decrease the failure rate drastically. German logistic companies report a first-time failure delivery rate of 4–8%, see [5].

3 Advanced Delivery Locations

In this section, alternative modern delivery concepts will be discussed. Referring to [15], the different types of parcel lockers can be categorized into:

- **Reception boxes (parcel boxes)**, which are permanently fixed at the customer's home. They can be accessed by a code or key.
- **Delivery boxes**, which are filled with goods at the depot of the logistics service provider and attached temporarily to the home with some locking device.
- **Controlled access systems**, which can be accessed by the delivery driver by a key or something similar.
- **Collection points**, like stores, post offices, or petrol stations, with usually long opening hours.
- **Locker-banks (parcel locker stations)**, which can be located near apartment blocks, work places, parking areas, or railway stations. This is a collection of several parcel boxes so that deliveries of various recipients can be stored and without the need of permanent recipient-specific boxes.

3.1 Parcel Locker Stations

The very first pilot project with parcel lockers was initiated by DHL in the year 2001 in Dortmund and Mainz in Germany. Since this time, the project became more and more common and today DHL operates 3,400 parcel lockers in Germany, which results in the possibility to reach a DHL parcel station in up to 10 min for 90% of the German citizen. Other similar projects are InPost lockers, U.S. Postal Service lockers, Post24 Parcel Machines and others.

Fig. 3. Exemplary DHL parcel locker in Nürnberg (from https://commons.wikimedia.org/wiki/File:BigPS.JPG)

The technological concept is always pretty similar. The parcel locker station consists of slots of differing sizes (from 15 cm × 11 cm × 1 cm up to 65 cm × 61 cm × 35 cm) and a central control panel, see e.g. Fig. 3. The deliverer needs to find an appropriate and empty slot to insert the parcel. The recipient is informed via eMail and/or SMS and receives a validation code for opening his parcel slot. For some companies, the recipient requires a membership card for validation. Additionally, it is possible that deliveries can be redirected to a parcel locker if the recipient was not at home. In such cases, the recipients receive a notification card which can be scanned at the panel to obtain the parcel.

Addressing parcel lockers can be done similar to usual home addresses. The customer needs to state the address and number of the parcel locker and his name and membership number. The logistics service provider will hence deliver the shipment to the parcel station and store them up to nine days.

Advantages of Parcel Lockers. One major advantage is the reduced handling effort for the logistics service provider. Deliveries to several customers can be committed without the need to drive to each home address. Additionally, failed deliveries will not occur reducing the effort for redirecting or second delivery attempts. Hence, the effort for carrying parcels to recipients which are not home, can be omitted. Consequently, the overall costs and amount of work of the logistics service providers can be reduced resulting in environmental benefits, like less delivery vehicles on the streets and a reduction of congestion and emissions.

One advantage for the recipient is that deliveries to parcel lockers can be handled faster than home deliveries as they are closer to the carriers depots and it is not required to drive to each home addresses. Secondly, the recipient might anticipate not being home when the parcel will be delivered. Hence, the parcel might be redirected to a neighbor which can have its drawbacks as well. The neighbor could not be home when the recipient reaches home, or the recipient itself gets home late and does not want to disturb any more. Besides from that, the parcel can be delivered by a second attempt, to a parcel locker, or a store of the logistics service provider such that the recipient will receive his shipment one or even more days later. If the customer on the other hand orders to a parcel locker, he can pick up his shipment at any time, e.g. after his working time on his way home, or at a parcel locker close to his daily shopping locations.

Disadvantages of Parcel Lockers. For most companies, membership cards are required which need to be ordered initially, is the initial hurdle that has to be taken by the customer. Another reason, which maybe has the biggest impact, is that the recipients need to actually decide over their delivery preference. In most scenarios, it is simply more straightforward to order parcels and shipments to their respective home addresses and hope for the best possible delivery (to home, neighbors, a near store, or in additional trials).

Last but not least it is not beneficial for the customers to order their delivery to a parcel locker if they anticipate that they will be at home during the estimated home delivery time.

When the effects of parcel locker deliveries are discussed, it is necessary to think about the resulting CO_2 emissions of the overall transportation. In the case without considering parcel lockers, the CO_2 emissions result from the vehicles delivery tour plus some possible additional delivery trials and re-deliveries towards stores or other drop points. If on the other hand, deliveries towards parcel lockers are offered, a huge amount of the delivery tour emissions vanish, since the very last part of the haul is done by the recipients themselves. This very last part of the haul is taken by the individual recipients themselves and highly diverse regarding the resulting CO_2 emissions since the amount of detour for that pick-up and the used transportation mode (by foot, bike, car, bus, tram, train) cause a wide range of possible CO_2 emissions, see e.g. [10].

From the logistics service providers perspective, the locations of parcel locker stations has to be planned initially and revised over time based on experiences and demand changes.

3.2 Parcel Boxes

The idea of separate parcel boxes at home is closely related to parcel locker stations. Since the parcel box will be placed outside the customer's home, the concept is pretty similar to the classical mailbox. The costs for the box have to be paid by the customer or a rent based on the provider. Depending on the provider of the parcel box the opportunity for use differ slightly.

The "DHL Paketkasten" (see Fig. 4) is available in different sizes and with possibility to mount to a wall or freestanding. It is a robust box, which is to be opened and closed only by DHL and the recipient. Thus, several parcels can be stored but this box cannot be used by any other logistics service provider. This is a strong disadvantage because for the majority of online orders, the recipient does not receive the possibility to decide over the logistics service provider.

Fig. 4. Exemplary DHL Paketkasten and Parcellock Box (from https://www.paketda.de/dhl/dhl-paketkasten-fuss.jpg and https://www.parcellock.de/)

"Parcellock" (see Fig. 4) is a system initiated by DPD, GLS, and Hermes, which will be mounted permanently as well. The locked box can by accessed by entering a TAN code. Those codes can be used by different logistics service providers and even local courier services or others.

Besides these, there are lightweight constructions like "Paketsafe", "PakSafe", "PaketButler" (see Fig. 5) and others which can be mounted safely but without installation work. They follow the idea that the box is unlocked initially. If a parcel arrives, the deliverer can insert it into the box and will lock the box which can be accessed by the recipient later. These systems can be accessed by any logistics service provider but the main disadvantage is, that only one delivery can be inserted until the recipient turns home to open the lock.

Delivery boxes differ because they are only temporarily locked to the wall and the delivery process changes. The recipient has to order his parcel to a Lockbox address where his parcel will be packed into a delivery box and redirected towards the recipients home address, where the delivery box will be locked with a respective mechanism. When the recipient turns home, he can detach the box from his wall and open it with a code or key. Empty boxes can be collected by a separate tour or the next delivery tour of the delivery box provider.

Fig. 5. Exemplary Paksafe, PaketButler, and PaketSafe Boxes (from https://www.paksafe.de/, https://www.paketbutler.com/, and https://paketsafe.net/)

Advantages of Parcel Boxes. The main advantage of parcel boxes is that the recipient can simply order deliveries to his home address. In a no-one-at-home scenario, the parcel does not need to be redirected, it can simply be inserted into the parcel box. Hence, there is no additional effort for the recipient to collect his parcel at all, he will either receive his parcel with the usual delivery attempt or in his parcel box, so there is no additional time, costs, and resulting emissions for his parcel collection and the logistics service provider.

For delivery boxes, there are only minor additional costs to receive a secure home delivery with no additional effort for the recipient.

Besides that, the logistics service provider obtains the advantage that there are no additional delivery trials or redirections. Hence, costs, time, effort, and the resulting environmental impacts decrease.

Disadvantages of Parcel Boxes. The main drawback is the resulting cost to provide the parcel box, which range from 100€ for the lighter systems up to 900€

for the more robust and taller boxes. For the lighter systems, monthly payments starting at 5€ are possible as well. On the other hand, the robust systems have to be mounted outside the house, which can be difficult if not even impossible to realize for most rental apartments in urban areas. The lighter boxes can be attached at or in front of the door, but they can only store one delivery due to the used locking system. Additionally, having a parcel box hanging on the apartment door might be unpleasant for some people.

The delivery box on the other hand does not capitalize on the straightforward home delivery scenario. The recipient has to order his parcel to the delivery box provider, who will redirect this packed parcel to the recipients address. Hence, the recipient needs to decide for each specific parcel if he prefers the usual home delivery or the delivery box detour with additional time, costs, and organizational effort for him and the delivery box provider.

3.3 Controlled Access Systems

Besides from the introduced different types of parcel lockers, one increasingly regarded concept is the possibility of delivering parcels to some type of controlled access systems like directly behind the front door, in a garage, or the trunk of the recipients car, see e.g. [24]. The logistics service provider will be authorized by the customers to access the location key-less and insert the respective delivery. Of course the location has to be equipped with a respective smart locking system. Different global and local acting companies and car manufacturer have become aware of this possibility and some early pilot projects are tested already. The arising benefits and drawback will be given below.

Advantages of Controlled Access Systems. Since the secure delivery locations offer enough space, there is no need to have additional boxes available in order to receive the parcel. Hence, there is no additional cost for the recipient nor the logistics service provider by using this opportunity. Besides that, the effort to collect the parcel from these systems is minor.

Disadvantages of Controlled Access Systems. Besides the main advantages, there are some drawbacks, which cannot be neglected. First of all, a modern infrastructure for the respective location has to be provided.

Regarding the trunk delivery, most cars are not fixed to one location during the whole day. Customers need to give different parking locations during the day which can be tracked by GPS additionally. Following [20], the arising task for the logistics service provider is a vehicle routing problem with roaming delivery locations with some certainty about the actual locations. Due to its complexity, it is worthwhile to consider problem specific solution approaches like the branch and price algorithm in [18].

Additionally from the present point of view, there is only a minor amount of actual implemented approaches which can be used by customers.

Finally, even though logistics service providers deal carefully with user data, security concerns have to be mentioned. The deliverer receives access to the private area of the recipient and in case of hacking attempts troubleshooting needs to react fast to secure customer's locking systems.

3.4 Summary of Advanced Delivery Locations

According to the given discussions about the respective advanced delivery locations Table 1 summarizes their main characteristics. The unique and recurring handling effort are stated for the respective concepts and separated for recipient's (R) and logistic service provider's (LSP) perspective.

The recipient's unique effort ranges from the order of simple membership cards, over building costs for parcel boxes, to costs for controlled private access systems. The latter vary highly depending on if cloud based locking systems are already existing or available for the regarded location and the value of the customer's consent for providing this access. The recipient's handling effort on the other hand refers to the effort to order and collect each specific parcel, which range the pick-up from a parcel locker station over the handling of delivery boxes to the simple collection from parcel out of individual boxes or controlled access systems.

From the logistic service provider's point of view, the unique effort arises from the need to plan, schedule and construct the respective infrastructure of the different concepts. This can be the need to actually build parcel locker stations after evaluating reasonable locations, the purchase of delivery boxes and systems, or the preparation of smart locking applications for the deliverer. The handling effort changes compose from the partially reduced transportation requirement,

Table 1. Evaluation of advanced delivery locations

	Parcel locker station	Parcel box	Delivery box	Front door or garage	Car trunk
R's unique effort	Minor	High	Minor	Depend	Depend
R's handling effort	Medium	Negligible	Minor	Negligible	Negligible
LSP's unique effort	High	None	Small	Minor	Minor
LSP's handling effort	High decrease	Small decrease	Small	Small	Medium
CO_2e reduction	High	Small	Small	Small	Small

the decreasing amount of failed deliveries, and the need to handle the respective locking system.

The CO_2e reduction potential arises from the decreasing amount of fail deliveries with its consequences. Only for the parcel locker stations, recipient specific deliveries disappear and hence the logistics service provider's handling effort as well as the overall emissions reduce considerably.

4 Advanced Delivery Vehicles

In this section, recent trends and future ideas for alternative delivery vehicles will be presented. There are various start-up concepts and pilot projects of different logistics service providers which try to overcome different issues of the last mile delivery. The majority of these concepts combines ideas of autonomous vehicles with local emission-free vehicles.

4.1 Electric Delivery Vehicles

For a recent review of past research and future perspectives of the opportunities see e.g. [19]. There are different ideas of electric vehicles in the parcel delivery, varying in vehicle type (trucks, vans, motorcycles, bikes, ...), vehicle size, and consequently in load capacity, driving range, and the appropriate use case. Since the scope of this work focuses on the last mile delivery, there are two main concepts which are very relevant and receive more and more application: electric mini vans and electric cargo bikes.

Fig. 6. Box-type, pick-up, and full size box van by StreetScooter (from http://cdn4.spiegel.de/images/image-1191617-860_galleryfree-cbhi-1191617.jpg, https://www.streetscooter.eu/, and https://insideevs.com/ford-dhl-unveil-streetscooter-work-xl-electric-truck/)

The development of electric delivery vans has increased drastically over the last five years. The most known and used concepts are the "Chanje Electric Delivery Vans" for FedEx, "Mercedes Electric Sprinter" and "Workhorse NGEN Electric Delivery Van" for different operators, and the "DHL StreetScooter", see Fig. 6. What all of these concepts have in common is that the range is between 80 km and 250 km, they drive fully electric, there load capacity is between 500 kg and 1000 kg, and they have limited top speed (90–120 km/h) to extend the available range and since their typical use case are city areas.

Advantages of Electric Delivery Vans. The typical advantages of electric vehicles can be transferred to electric delivery vans with some additional remarks. Of course, they drive local emission-free which leads to cleaner cities. Especially regarding the current trend of low-emission zones in more and more cities worldwide, electric delivery vans will have a great potential. Besides that, DHL reports to save 16,000 tons of CO_2 per year and 60 to 80% costs for maintenance and wear of the vehicles.

Disadvantages of Electric Delivery Vans. The major disadvantage of electric delivery vehicles is again the high initial costs for vehicles and infrastructure. At the moment, prices for StreetScooters range from 40,000€ up to 55,000€ depending on type, size, and battery capacity. Additionally, for the most companies it is beneficial to provide own charging infrastructure, which has to be build and maintained as well. The limited driving range has to be mentioned as well, since it is not practical to use these types of vehicles on mid or long haul delivery tours.

4.2 Electric Cargo Bicycles

Besides vans, usually bicycles are used to delivery letters and small parcels like books or brochures. [4] and [13] describe the biggest challenges and chances by using electric cargo bikes in the delivery process. Since mail carriers might drive up to 25 km per day and with 50 kg, companies started using electric bicycles more and more. Deutsche Post reports to use approximately 13,000 electric bikes for this purpose, see [8].

Fig. 7. Parcylces and Cubicycle of DHL (from https://www.electrive.net/2017/05/30/dhl-express-schickt-cubicycle-nach-berlin/)

Additionally, first pilots have started to delivery even parcels by appropriate electric bikes. So called cargo-bikes and -trikes have been developed by different manufacturers which allow to transport more heavy and bulky parcels. The largest electric cargo bike fleet so far is operated by DHL. They use a combination of electric bicycles, and electric cargo bikes, which they call “Parcycle” and

"Cubicyle" depending on the type and size, to manage different parcel sizes and weights, see Fig. 7. The Cubicycle can carry up to 125 kg. The cargo boxes can be mounted and unmounted on a railing system of Cubicycles from a trailer to deliver several parcels with bike tours, see Fig. 8. If the trailer is located at the depot or temporarily unmoved at any other locations, the tours can be regarded as closed tours. On the other hand, it is possible to relocate the trailer during the bike tours, so that the bike tours become open tours. Hence, the need to synchronize the routes of both vehicle types arises whereby it is possible to handle larger areas more efficiently. Pilot projects with Parcycles and Cubicyle have been initiated in Europe, mainly in the Netherlands and Germany, see [7].

Fig. 8. Mounting cargo boxes on Cubicycles of DHL (from https://www.velove.se/news/city-containers-new-pilot-dhl-express-frankfurt-utrecht)

Advantages of Electric Cargo Bicycles. Electric cargo bikes enable to utilize certain advantages of bicycle traffic in opposite to delivery vans. From energy perspective, they do not use classic fuels but rechargeable batteries so that they drive local emission-free and much more economically. If available, cargo bicycles are allowed to drive on bikeways, which are usually way less congested than roads. Hence, the driver will be able to work less stressful and in certain scenarios in big cities even faster than classical van deliveries. Apart from that, parking bicycles requires way less space which is positive for all traffic participants and citizen. DHL deliverers report a very positive experience using cargo bicycles based on these positive effects.

Disadvantages of Electric Cargo Bicycles. The drawbacks of cargo bikes are negligible. Of course, initially electric cargo bikes have to be bought and a respective charging infrastructure has to be provided. The typical load capacity (for weight and volume) of all types of cargo bikes is way less compared to classic delivery vans. For these large and heavy parcels, classic delivery vans or even hauliers will still be required. Regarding the route scheduling, using mountable cargo boxes becomes more complex than classic delivery van routes since it is necessary to manage trailer locations(s), delivery sorting into separate cargo boxes, (synchronized) cargo bike routes, and the required deliverers.

4.3 Autonomous Delivery Vehicles

Since several years, logistic service providers, start-ups, and researchers have shown potentials and prototypes of autonomous delivery vehicles of various types and sizes, see e.g. [12] and [17]. There are technological assistants, like adaptive cruise control, lane, break, blind spot, parking, or hill-holder assistants, but what this section will concentrate on are unmanned, Level-5 self-driving vehicles, like small cars, drones, or robots because they allow to reduce the working effort for deliveries. There are projects where autonomous, mostly electric, robots are used in intra logistics but the usage in road traffic remains a theoretical idea mostly. Two of the few actual implemented unmanned concepts are the electric delivery vehicle R1 by Nuro and Kroger (see [21]) and Starships's mobile robots (see [22]), displayed Fig. 9.

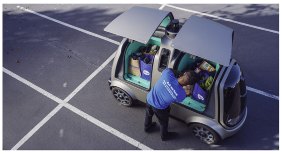


Fig. 9. Nuro's R1 vehicle and Starship robot (from https://spaces.hightail.com/space/hFlqnnpX76/files and https://www.starship.xyz/)

The R1 is used as a part of the vehicle fleet to fulfill customer grocery orders by Fry's Food stores in Arizona. Starship's robots are tested in above 100 communities and 20 countries worldwide and move on walkways in walking speed.

Advantages of Autonomous Delivery Vehicles. The big chance of autonomous delivery vehicles is that they combine the positive aspects of electric vehicles and sacrifice drivers. Hence, this manpower can be used to stressless manage the vehicles and organizational tasks of the delivery process.

Disadvantages of Autonomous Delivery Vehicles. Since fully autonomous vehicles move unmanned, several countries still do not have appropriate laws concerning the actual conditions for driving permissions on roads, bikeways, walkways, etc. This is additionally based on the fact that there are still doubts that autonomous vehicles can safely manage unpredictable and unusual situations. Especially on walkways, robots need to deal with diverse road conditions. Again, adding autonomous delivery vehicles into an existing fleet is costly and the amount of manufacturers of appropriate vehicles is still small.

5 Conclusion and Further Concepts

Besides the described innovations, a variety other concepts to improve the last mile delivery exists. Amazon Go [1] and the Chinese start-up BingoBox [27] are the best known examples for staff free small supermarkets. A combination of an indispensable smartphone app, RFID chips, and several cameras and sensors allows customers to walk in, pick or replace any products, and simply leave the store without the need to checkout or scan the products at all. The payment is done with the help of the respective smartphone app and hence the customer saves a lot of time.

Certain concepts for autonomous vehicles, drones, or robots have been studied, like the Mercedes Vision Van which combines an electric van with a railing system to exchange the complete cargo bay including the load. Additionally, drones are attached to the van assisting the delivery in cities. Other drone delivery concepts are investigated by various companies, see e.g. [2]. Unfortunately, most of these concepts have not made it into real world pilot projects so far.

What this work reveals are many different approaches and concepts tackling the main drawbacks in last mile delivery. These concepts are presented separately, but in real world applications they can be combined strategically. The aim of logistic service providers should be to reduce their costs and transport emissions and still maintain or even increase customer satisfaction. Shifting as much deliveries as possible onto electric vans or cargo bikes is one very worthwhile option. Besides that, consolidating deliveries has to be observed more detailed, since home address deliveries are the most costly part for several reasons. Logistic service providers could analyze the density of deliveries among certain areas and usual commutes to place parcel locker stations strategically. Of course, this parcel locker station network needs to be rescheduled over time based on experiences and customer behaviors. An operative approach should be to consider service provider independent parcel locker stations or provide access to already existing systems, which would allow several big and even small companies to access these lockers. Of course, an appropriate security system as well as cost distribution among the logistics service providers has to be specified. These and other recent and future trends are worthwhile approaches in order to reduce environmental effects and costs of logistic service providers and increase the benefit of the customers.

Acknowledgement. This research is funded by the German Research Foundation (DFG) under reference BI 555/4-1.

References

1. Amazon. Amazon Go (2019). https://www.amazon.com/b?node=16008589011. Accessed 18 Mar 2019
2. Amazon. Amazon Prime Air (2019). https://www.amazon.com/Amazon-Prime-Air/b?node=8037720011. Accessed 18 Mar 2019

3. Amazon. Prime FREE Same-Day and One-Day (2019). https://www.amazon.com/b?ie=UTF8&node=8729023011. Accessed 18 Mar 2019
4. Anderluh, A., Hemmelmayr, V.C., Nolz, P.C.: Chapter 8 - Sustainable logistics with cargo bikes—methods and applications. In: Faulin, J., Grasman, S.E., Juan, A.A., Hirsch, P. (eds.) Sustainable Transportation and Smart Logistics, pp. 207–232. Elseivier (2019)
5. Bundesverband Paket und Expresslogistik e. V. (BIEK). Nachhaltigkeitsstudie (2017). https://www.biek.de/download.html?getfile=508. Accessed 18 Mar 2019
6. Bundesverband Paket und Expresslogistik e. V. (BIEK). KEP-Studie 2018 - Analyse des Marktes in Deutschland (2018). https://www.biek.de/download.html?getfile=1928. Accessed 18 Mar 2019
7. Deutsche Post DHL Group. DHL expands green urban delivery with city hub for cargo bicycles (2017). http://www.dhl.com/en/press/releases/releases_2017/all/express/dhl_expands_green_urban_delivery_with_city_hub_for_cargo_bicycles.html. Accessed 18 Mar 2019
8. Deutsche Post DHL Group. Deutsche Post nimmt 1000. E-Bike von StreetScooter in Betrieb (2019). https://www.dpdhl.com/de/presse/pressemitteilungen/2017/deutsche-post-puts-1000th-e-bike-from-streetscooter-into-service.html. Accessed 18 Mar 2019
9. Deutsche Post DHL Group. Same day (2019). https://www.dhl.de/en/express/export-services/same-day.html. Accessed 18 Mar 2019
10. Edwards, J., McKinnon, A., Cherrett, T., McLeod, F., Song, L.: The impact of failed home deliveries on carbon emissions: are collection/delivery points environmentally-friendly alternatives? In: Annual Logistics Research Network Conference, vol. 14, pp. 102–108 (2009)
11. FedEx Corporation. Same Day Delivery—FedEx (2019). https://www.fedex.com/en-us/shipping/same-day.html. Accessed 18 Mar 2019
12. Figliozzi, M.A.: Lifecycle modeling and assessment of unmanned aerial vehicles (drones) CO_2e emissions. Transp. Res. Part D: Transp. Environ. **57**, 251–261 (2017)
13. Gruber, J., Kihm, A.: Reject or embrace? messengers and electric cargo bikes. Transp. Res. Procedia **12**, 900–910 (2016). Tenth International Conference on City Logistics 17–19 June 2015, Tenerife, Spain
14. Handelsverband Deutschland e.V. (HDE). Zahlenspiegel 2018 (2018). https://einzelhandel.de/index.php?option=com_attachments&task=download&id=9593. Accessed 18 Mar 2019
15. Iwan, S., Kijewska, K., Lemke, J.: Analysis of parcel lockers' efficiency as the last mile delivery solution - the results of the research in Poland. Transp. Res. Procedia **12**, 644–655 (2016). Tenth International Conference on City Logistics 17–19 June 2015, Tenerife, Spain
16. Kopun, F.: Online shopping drives surge in holiday returns, and what happens next may surprise you (2017). https://www.thestar.com/business/2017/01/15/online-shopping-drives-surge-in-holiday-returns-and-what-happens-next-may-surprise-you.html. Accessed 18 Mar 2019
17. Mahmassani, H.S.: 50th anniversary invited article - autonomous vehicles and connected vehicle systems: flow and operations considerations. Transp. Sci. **50**(4), 1140–1162 (2016)
18. Ozbaygin, G., Karasan, O.E., Savelsbergh, M., Yaman, H.: A branch-and-price algorithm for the vehicle routing problem with roaming delivery locations. Transp. Res. Part B: Methodol. **100**, 115–137 (2017)

19. Pelletier, S., Jabali, O., Laporte, G.: 50th anniversary invited article - goods distribution with electric vehicles: review and research perspectives. Transp. Sci. **50**(1), 3–22 (2016)
20. Reyes, D., Savelsbergh, M., Toriello, A.: Vehicle routing with roaming delivery locations. Transp. Res. Part C: Emerg. Technol. **80**, 71–91 (2017)
21. Robotics Business Review. Nuro, Kroger add Driverless Vehicle R1 to Unmanned Grocery Delivery Fleet (2018). https://www.roboticsbusinessreview.com/unmanned/nuro-kroger-unmanned-grocery-delivery-fleet/. Accessed 18 Mar 2019
22. Robotics Business Review. Starship Launches On-Demand Mobile Robot Package Delivery (2018). https://www.roboticsbusinessreview.com/news/package-delivery-starship-launches/. Accessed 18 Mar 2019
23. Saleh, K.: E-commerce product return rate - statistics and trends (2017). https://www.invespcro.com/blog/ecommerce-product-return-rate-statistics/. Accessed 18 Mar 2019
24. Savelsbergh, M., Van Woensel, T.: 50th anniversary invited article - city logistics: challenges and opportunities. Transp. Sci. **50**(2), 579–590 (2016)
25. Smith, E.G.: How online retailers predict your perfect outfit (2015). https://www.wsj.com/articles/how-online-retailers-predict-your-perfect-outfit-1438794462. Accessed 18 Mar 2019
26. Speranza, M.G.: Trends in transportation and logistics. Eur. J. Oper. Res. **264**(3), 830–836 (2018)
27. Tech in Asia. In China, Amazon's 'store of the future' is already open (2017). https://www.techinasia.com/china-version-amazon-go-bingobox-funding. Accessed 18 Mar 2019
28. Weltevreden, J., Mindali, O.R.: Mobility effects of b2c and c2c e-commerce in the netherlands: a quantitative assessment. J. Transp. Geogr. **17**(2), 83–92 (2009)

Reconfiguration of the Last Mile: Consumer Acceptance of Alternative Delivery Concepts

Vanessa Felch, David Karl(✉), Björn Asdecker, Amelie Niedermaier, and Eric Sucky

University of Bamberg, Feldkirchenstraße 21, 96050 Bamberg, Germany
david.karl@uni-bamberg.de

Abstract. In the coming years, the growth of e-commerce is expected to continue. Due to the increasing number of parcels delivered and existing capacity bottlenecks within the logistics service providers, concepts are emerging that are newer and more sustainable than traditional home delivery. This article analyzes (1) how customers perceive four alternative concepts (reception box, controlled access systems, trunk delivery, and crowdsourced delivery among friends) and (2) what influences acceptance, which is the decisive prerequisite for continued application. The data of 207 young German consumers were analyzed using descriptive methods and linear regression. The results show that the reception box achieves the highest acceptance level. In addition, perceived usefulness, security, and privacy are crucial factors in the customers' intention to use delivery concepts. Since the study provides indications of the concepts' success potential, it is of value to decision-makers from e-tailers, logistics service providers, and politicians that intend to incorporate new delivery concepts.

Keywords: Last mile · Alternative delivery concepts · E-commerce · Consumer acceptance · Crowdsourced delivery · Controlled access systems · Trunk delivery · Reception box

1 Introduction

While global e-commerce revenues reached a new record of $2.3 trillion worldwide in 2018, this figure is expected to double by 2021 (eMarketer 2019). The continued growth will be accompanied by an increasing number of parcels and thus an increase in deliveries (Reyes et al. 2017; Marujo et al. 2018). Within the delivery chain, the so-called "last mile", that is, the physical distribution of parcels between the last transit point (e.g., the hub of a logistics service provider) and the final drop-off point, is the biggest challenge (Morganti et al. 2014; Wang et al. 2014, 2016; Reyes et al. 2017; Devari et al. 2017). Due to the "[…] trade-offs between routing efficiency and consumer convenience" (Kull et al. 2007, p. 441) the last mile constitutes a bottleneck and accounts for up to 30% of the total distribution costs (Wang et al. 2014, 2016).

On the one hand, the high costs can be attributed to a small number of successful first delivery attempts (Wang et al. 2016; Macharis and Melo 2011). On the other hand, previously agreed-upon time window deliveries facilitate inefficiencies (Macharis and

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 157–171, 2019.
https://doi.org/10.1007/978-3-030-29821-0_11

Melo 2011; Punakivi et al. 2001). Time windows lead to longer routes to deliver the same number of parcels. In addition, both issues have an impact on the three sustainability dimensions, namely, economic, environmental, and social aspects (Devari et al. 2017; Macharis and Melo 2011). These include but are not limited to more traffic congestion and wasted resources, increased emissions and air pollution, stress, and unfavorable working conditions (Macharis and Melo 2011; Anderson et al. 2005). Consequently, Gevaers et al. (2014) concluded that the last mile delivery is "[…] currently regarded as one of the more expensive, least efficient and most polluting sections of the entire logistics chain" (p. 56).

Against this background, both businesses and academia have suggested new, more sustainable delivery concepts that reduce the inefficiencies of the traditional home delivery. For instance, a couple of years ago, DHL introduced the "Packstation", a publically available reception box. Now this concept has been scaled down to private mailboxes for parcels. In addition, the largest e-commerce retailer, Amazon, gained attention with innovative concepts such as Amazon Key (delivery via a controlled door access system) and thus intensified competition for consumer interest (Amazon 2018, 2017). In addition to Amazon's innovations, collaborations are emerging between automobile manufacturers and logistics service providers, who are jointly working on trunk deliveries (Volvo 2016; Audi 2015). From the academic side, Devari et al. (2017) suggested adapting the idea of crowdshipping to delivery among friends. Therefore, while there are many conceptual proposals for increasing efficiency in the last mile, it is still unclear how consumers perceive these new concepts and what influences acceptance as the crucial prerequisite for continued application. Småros et al. (2000) argued that market success depends more on consumer acceptance than on technical and technological solutions for problems. For this reason, this work pursues the following two research questions:

- RQ1: What are consumers' perceptions of the currently discussed alternative delivery concepts (ADCs) for the last mile?
- RQ2: How do different factors influence consumer acceptance of ADCs?

This paper contributes to the last mile design/configuration literature (e.g., Anderson et al. 2005; Hayel et al. 2016; Gevaers et al. 2014). The publication that is most similar to this one is from Yuen et al. (2018), who investigated the intention of consumers to utilize self-collection services. By using Innovation Diffusion Theory (IDT), Yuen et al. were able to identify attributes that influence the decision to use an innovation. In total, they surveyed 164 consumers in Singapore and analyzed the data using regression analysis. The presented model enabled them to explain 42% of the variance in consumers' intention to use self-collection services. In their conclusion, Yuen et al. (2018) point out that "[…] there are other factors that can influence consumers' intention" (p. 7) and these require additional behavioral studies. They also highlight that other concepts should be examined in follow-up studies. This paper responds to this call for research and investigates alternative home delivery concepts instead of self-collection services and considers additional influencing factors.

2 Theoretical Background

2.1 Alternative Delivery Concepts for the Last Mile

Many retailers and courier, express and parcel service providers seek to offer fair and responsible logistics services. Therefore, they have a need for more sustainable delivery concepts to distribute their products more efficiently (DHL 2016). Two different delivery methods can be distinguished: (1) consumer pickup and (2) home delivery (Hayel et al. 2016). In the case of self-collection, parcels are delivered to either fully automated lockers, such as the DHL Packstation, or a parcel shop close to home or work, where consumers pick up the parcel themselves (Yuen et al. 2018). Thus, consumers cover the last mile themselves (Hayel et al. 2016). Studies by Tan (2016) and Morganti et al. (2014) in Singapore and France highlight that only up to 20% of the participants prefer consumer pickup to home delivery. A survey conducted in Germany provides similar results (PWC 2018). Only about half of the respondents could imagine delivery to a parcel shop or to fully automated lockers under the condition that opening hours and availability are improved.

On the other hand, home delivery can be divided into two types: (1) attended and (2) unattended (Yuen et al. 2018). In the former case, service providers hand parcels over to consumers, who usually need to be present during delivery (Xu et al. 2008). This leads to inefficiencies, as described in the introduction. The situation is different with unattended delivery. Originally, the parcels were left on the stairs of the house or in the garden (Xu et al. 2008). The transfer of the parcel to the consumer's area of responsibility irrespective of the consignee's presence characterizes this delivery type. This allows minimizing the number of failed deliveries to a single successful attempt, reduces redundancies and thus increases the efficiency of the last mile (Xu et al. 2008; Punakivi et al. 2001). Due to various security concerns, especially with high parcel values, new solutions with access to private properties, such as controlled access systems or reception boxes, have emerged (Xu et al. 2008).

This study takes a closer look at these unattended delivery concepts in which the parcel is not necessarily delivered directly to the consumer's home but is brought into their area of responsibility. Thereby, we will focus on those concepts recently proposed by businesses and/or discussed in the literature: (1) reception box, (2) controlled access systems, (3) trunk delivery, and (4) crowdsourced delivery among friends (Nabot and Omar 2016; Iwan et al. 2016; Capgemini 2019; Faulin et al. 2019; Yuen et al. 2018; Devari et al. 2017). Other concepts, such as delivery to neighbors, are not taken into account, as they constitute a favor and are not a systematic, reliable delivery option. Furthermore, alternative propulsion concepts, such as natural gas or hydrogen, e-scooters, and autonomous delivery vehicles are excluded from the study, since they do not solve the actual capacity problem on the last mile.

Reception Box

The reception box, often referred to as "mailbox for parcels", is a fixed, consumer-specific box for receiving and returning parcels (Punakivi et al. 2001). The box is permanently attached to the front of the house or placed on the ground in front of it (Iwan et al. 2016). During parcel delivery, the service provider opens the reception box using a key or an electronic variant such as a single-use code (Wang et al. 2014; McKinnon and Tallam 2003). Once the box has been closed, the system generates a message informing the consignee about the successful delivery. From a consumer perspective, the incompatibility of provider-specific reception boxes, such as the DHL parcel box, is a crucial disadvantage. In contrast, supplier-independent solutions allow different service providers to use the same reception box for delivery.

Controlled Access Systems

In contrast, the integration of controlled access systems represents a smart home solution with access to the inside of the apartment or house. This concept is currently offered by Amazon with its Amazon Key or by Walmart in cooperation with August (Amazon 2017; Walmart 2017). Through an encrypted authentication process, during which the address and planned delivery time is compared, the service provider receives access to the apartment or house (Iwan et al. 2016; McKinnon and Tallam 2003). Consumers can track their delivery in real time or retrospectively. As soon as the delivery has taken place, a notification is sent to the consumer.

Trunk Delivery

In addition to an apartment, a car trunk can also serve as a mobile delivery point for parcels. The consignee's car is parked at a specified delivery address for a predefined period of time (Ozbaygin et al. 2017; Mandziuk 2019). The service provider locates the car and opens the trunk with a single-use digital key (Reyes et al. 2017). Parcels can be delivered and returned without the presence of the consignee. To take advantage of this concept, a connectivity box must be retrofitted in the trunk of the vehicle (Robarts 2016). So far, the automobile manufacturers Volvo, Volkswagen, and Audi have cooperated with Amazon to provide an in-car delivery service (Volvo 2016; Audi 2015; Reyes et al. 2017). In addition, DHL and Smart have launched a pilot project in Germany (Robarts 2016).

Crowdsourced Delivery among Friends

On-demand transportation companies, such as Uber and Lyft, have significantly changed the market (Glöss et al. 2016). The idea of outsourcing services to the anonymous masses has led to the development of new business models. Logistics also draws on the principle of crowdsourcing. Parcel deliveries are outsourced to the anonymous masses, mostly private drivers, who are financially remunerated (Wang et al. 2016; Castillo et al. 2017). However, a lack of trust, reliability, and privacy are the biggest problems of crowdsourced delivery (Devari et al. 2017). Therefore, Devari et al. (2017) developed the concept further. In their approach, friends of the consignee serve as the delivery party. The process is handled through a platform that uses data from social networks. To coordinate delivery, existing, mutually known communication channels can be used. For instance, friends from the same residential area can take parcels with them on their way home. This means that delivery routes, and thus emissions, can be reduced.

In terms of consignee presence, all four concepts are more flexible than attended home delivery. The first delivery attempt is successful; further attempts are not required, so the process becomes less time-consuming. The alternative concepts offer many advantages to consumers, logistics service providers, e-tailers, and the environment. Therefore, the concepts are empirically analyzed with regard to consumer acceptance.

2.2 Hypothesis Development

As an early indicator of an ADC's general consumer acceptance, we investigated influences on the willingness to test an ADC for free. This approach differs from Yuen et al. (2018), who considered trialability as a predictor and the consumer's intention to use the ADC as the model's dependent variable. The IDT provides an appropriate basis to better understand how innovations spread. Of particular interest is the individual decision-making process on whether to accept an innovation, which, according to Rogers (1962), can be divided into five phases: (1) knowledge, (2) persuasion, (3) decision, (4) implementation, and (5) confirmation. During the decision phase, individuals weigh the benefits and risks of using the innovation and decide whether to accept or reject it. To identify various benefit and risk characteristics, we have synthesized the existing ADC literature and have used selected dimensions of established behavioral theories that explain the acceptance and use of innovations and information technology, namely, the Technology Acceptance Model (TAM), the extended TAM, the Theory of Reasoned Action (TRA), and the Unified Theory of Acceptance and Use of Technology (UTAUT). The assumed relationships were then translated into hypotheses.

Perceived Benefits of an ADC

On the one hand, an ADC needs to offer a performance that promises a "functional benefit" (Ramadan et al. 2017) to the consumer that generates "relative advantage" in contrast to the status quo (Rogers 1962; Moore and Benbasat 1991). This can be summed up as the "performance expectancy" (Venkatesh et al. 2003) of consumers. To achieve this benefit, a delivery concept must accomplish higher usefulness, flexibility, or punctuality. Higher delivery flexibility reduces the consumers' effort if the delivery time better fits their personal schedules or if they do not need to pick the parcels up at a store/post office after several failed delivery attempts. These factors affect the "effort expectancy" (Venkatesh et al. 2003), which describes the perceived ease of using an innovation. If a shipment is delayed or cannot be scheduled according to one's personal needs, the "perceived usefulness" (Davis et al. 1989) of the applied delivery concept will obviously be low. Consequently, punctuality and flexibility are central factors in the delivery performance. Punel and Stathopoulos (2017) and Mladenow et al. (2016) also mentioned punctuality as a critical factor for crowdsourced delivery, which can also be assumed for other ADCs. Drawing on the TRA, it can be concluded that if an individual believes that a certain behavior will lead to a desirable outcome, then he or she is more likely to have a positive attitude towards the behavior (Ajzen and Fishbein 1980). Accordingly, the expectation of a positive outcome of a delivery concept leads to the development of a positive attitude towards the concept, implying a higher degree of intention to test this concept. Based on this reasoning, we hypothesize the following:

H1. Perceived usefulness has a positive effect on consumers' intention to test an ADC.

H2. Perceived punctuality has a positive effect on consumers' intention to test an ADC.

H3. Perceived flexibility has a positive effect on consumers' intention to test an ADC.

Perceived Risks of an ADC

On the other hand, every innovation contains uncertainties, which influence the consumers' perceived risk (e.g., Dowling and Staelin 1994). The perceived risk of an innovation negatively influences the intention to use it (Pavlou 2003; Martins et al. 2014). Risks include potential personal harm, e.g., safety issues and the violation of privacy (Ramadan et al. 2017). Thus, the protection of privacy is crucial for the acceptance of a delivery concept, especially when private property is involved (e.g., the ability of the delivery service to open the entry door or car trunk). In accordance with drone delivery acceptance (Ramadan et al. 2017), security, the reliability of undamaged shipment success, and liability for damaged or lost shipments are important for the attitude towards a delivery concept and subsequently for its acceptance. The last two factors are similarly of essential interest in crowdsourced delivery (Devari et al. 2017), where responsibilities are difficult to determine. Park et al. (2016) accordingly classify the reliability of a delivery as essential. Therefore, the following hypotheses are proposed:

H4. Perceived security has a positive effect on consumers' intention to test an ADC.
H5. Perceived privacy has a positive effect on consumers' intention to test an ADC.
H6. Perceived reliability has a positive effect on consumers' intention to test an ADC.
H7. Perceived liability has a positive effect on consumers' intention to test an ADC.

3 Empirical Study on Consumer Acceptance of Alternative Delivery Concepts

With regard to the research questions, we pursue a quantitative empirical research approach. An online questionnaire was initiated to identify consumers' acceptance of the four concepts and factors that are essential for the acceptance of ADCs. The survey consisted of three parts. The first part collected data concerning the respondents' purchasing behaviors. The second part focused the evaluation of the four ADCs. Initially, the ADCs were briefly explained before respondents were asked to evaluate items regarding them on a five-point Likert scale (–2 "strongly disagree" to +2 "strongly agree"). In case of missing expertise, the respondents were allowed to choose "not sure". Based on the hypotheses, nine items were formulated. According to Ramadan et al. (2017), these items can be summed up in two dimensions: functional benefit and risk (Table 1). To allow for consistency checks that ensure data quality, the items of "security" and "usefulness" were additionally inverted-scaled. A total of 42 participants with contradictory answers had to be excluded from further analysis. Finally, the third part consisted of questions on demographic data (e.g., age, sex). A pretest by five experts led to minor changes regarding wording and design. As Germany is one of the largest European e-commerce markets alongside France and

Great Britain (Morganti et al. 2014), the study was distributed to various German universities and online platforms. The survey mainly involved young consumers aged 16–44, who constitute the majority of e-commerce consumers in Germany (Statistisches Bundesamt 2019b). The field phase took place in late December 2017 to January 2018. In total, 207 valid data sets were tested for potential late-response biases. On a level of α = .05, no significant differences were tested.

Table 1. Overview of the dimensions, items and statements used

Dimension	Item	Statement to be evaluated
Benefit	Usefulness	The delivery concept is useful.
	Punctuality	The delivery concept leads to a punctual delivery.
	Flexibility	The delivery is flexible with regard to the delivery day, the delivery time and my presence.
Risk	Security	The delivery concept is secure.
	Privacy	The delivery concept maintains my privacy.
	Reliability	The delivery concept and the supplier are reliable.
	Liability	With this delivery concept, it is obvious who is liable in case of parcel loss.
Willingness to test	Excluding acquisition costs	I would be willing to test the delivery concept without paying for the acquisition costs.
	Including acquisition costs	I would be willing to test the delivery concept if I had to pay the acquisition costs myself.

3.1 Characteristics of the Sample

Overall, participants were predominantly women (67%). As we observed significant differences in the responses of men and women, weights that reflect the actual population characteristics of Germany (Statistisches Bundesamt 2019a: 50.65% women, 49.35% men) are used for the subsequent analysis. The majority of the sample were students (67%); relatedly, the average age of the participants was 27 years. The most frequently ordered items were fashion items (77%), followed by electrical appliances (40%), and computer and equipment (28%). For online orders, the participants primarily chose Amazon (93%) and Zalando (36%). At the same time of the survey, none of the participants had experience with ADCs on the last mile. Therefore, the study investigates their willingness to test the ADCs.

3.2 Descriptive Results

The greatest usefulness was perceived by participants using the reception box, followed by controlled access systems (Table 2). Trunk delivery and crowdsourced delivery among friends also achieved positive average values in terms of usefulness. Deliveries

Table 2. Descriptive results of the survey (Mean / S.D. / n)

	Reception box	Controlled access systems	Trunk delivery	Crowdsourced delivery among friends
Usefulness	1.39 / .89 / 206	−.03/ 1.24 / 201	.14 / 1.21 / 204	.12 / 1.25 / 203
Punctuality	1.38 / .79 / 203	1.24 / .81 / 198	.79 / .87 / 203	−.64 / .97 / 200
Flexibility	1.74 / .62 / 205	1.47 / .81 / 202	.52 / 1.21 / 205	.37 / 1.23 / 204
Security	1.50 / .75 / 205	−.92 / 1.06 / 202	−.09 / 1.14 / 204	.33 / 1.17 / 205
Privacy	1.77 / .60 / 205	−1.40 / .91 / 206	−.07 / 1.24 / 206	.29 / 1.28 / 205
Reliability	1.55 / .62 / 202	.14 / 1.19 / 188	.47 / .99 / 193	−.01 / .97 / 194
Liability	.87 / 1.08 / 201	.38 / 1.29 / 197	.13 / 1.21 / 200	−1.08 / .97 / 199
Willingness to test (no costs)	1.59 / .85 / 205	−.69 / 1.43 / 202	.03 / 1.51 / 200	.23 / 1.41 / 202
Willingness to test (with costs)	−.27 / 1.32 / 205	−1.57 / .88 / 205	−1.31 / 1.06 / 205	.23 / 1.41 / 202

with access to private properties were regarded as more punctual. In contrast, more than half of the respondents (62%) believed that crowdsourced delivery among friends would not lead to on-time delivery. The flexibility of trunk delivery and crowdsourced delivery among friends was rated lower when directly compared to the two other ADCs.

Participants considered delivery to a reception box to be the most secure. Next to the other outcomes, the concept of controlled access systems was not considered to be secure. A similar result can be observed with regard to privacy. Overall, 97% of the respondents did not expect any privacy issues when receiving a parcel through a reception box, while more than 86% perceived delivery via a controlled access system to be too invasive. In addition, the reception box was rated as the most reliable concept, whereas the other concepts were assessed significantly lower. Almost 80% of all respondents had liability concerns and feared the loss of their parcel if friends were to take part in the delivery process. Additionally, in case of the controlled access systems and trunk delivery, the participants were unsure about a clear liability transition. Finally, the willingness to test the ADCs was investigated. On average, participants were much more willing to test the ADCs if the test phase was free of charge (Mean = .26, S.D. = 1.57) than if they would have to pay for it (Mean = −.74; S.D. = 1.39). This difference (BCa 95% CI: [.91, 1.09]) was significant (t (801) = 22.20, $p < .001$) and represented a large-sized effect (d = .78) according to Cohen (1988). Therefore, it can be stated that consumers' intention to test ADCs is significantly higher if the expenses are paid by the logistics service provider.

In general, participants rated the reception box above average in the dimension benefit when the dimension risk was sufficiently considered. Compared to other concepts, the reception box was perceived as the most flexible and punctual alternative with the greatest usefulness. With regard to these three items, the controlled access systems received a similarly high rating. However, respondents have far greater risk concerns about controlled access systems than they do about reception boxes. Hence, it can be assumed that the consumer acceptance of the reception box is higher than for any other examined concept. Crowdsourced delivery among friends and trunk delivery received a medium level of acceptance. The controlled access systems received a medium rating for functional benefit and a low rating in terms of perceived risk. The following section offers more details how the previously discussed factors influence consumer acceptance of ADCs.

3.3 Regression Analysis: Influencing Factors on Consumer Acceptance of ADCs

Consumer acceptance of ADCs is operationalized by the consumers' willingness to test one of the concepts with all expenses for the test phase being paid by the logistics service provider. This intention is regressed on the predictors in two steps to uncover the impact of the control variables compared to the hypothesized influence factors. Model 1 includes all consumer-sided control variables and the delivery concept that is evaluated by the participants. Model 2 uses the full set of items that are supposed to influence the operational intent to test the delivery concepts. For every participant, four records (one for each ADC) can be used. After a listwise deletion of incomplete cases, 675 of theoretically possible 828 cases constitute the dataset. The regression analysis was performed by SPSS 24. Table 3 presents the model results from the regression analysis that will be discussed afterwards.

Model 1 contains several demographic characteristics (age, gender, education, monthly income, rural vs. urban residential area) of the respondents as well as information on their online shopping behavior (frequency, Amazon Prime membership). Each case was assigned to the specific ADC for which the survey questions were designed. The reception box was chosen as reference category while controlled access systems, trunk delivery, and crowdsourced delivery among friends were modeled by binary variables in each case. The model ($R^2 = .320$, $p < .001$) is significant. The variance inflation factors (VIF) are between 1.04 and 1.47. Thus, multicollinearity is not a major concern for this model, as it is far below a critical threshold of 10 (Field 2013).

Model 2 extends the first model by integrating the suspected factors for the consumers' acceptance, namely, the perceived security, privacy, reliability, liability, usefulness, punctuality, and flexibility. The model ($R^2 = .625$, $p < .001$) and the inclusion of the additional variables ($\Delta F = 76.533$; $p < .001$) is highly significant. Compared to Model 1, the variance explained is increased by 30% points. Thereby, the full model explains 62.5% of the variance. As the VIF of the independent variables are 3.59 at most, there is no obvious multicollinearity problem. The value of the Durbin-Watson statistic is very close to the value of 2, which indicates no autocorrelation (Field 2013). Because we identified some heteroscedasticity (HC) problems in the method according to Glesjer (1969), we ran a heteroscedasticity-robust regression according to the proposal of Hayes

Table 3. Results of the OLS regression models

	Mean	Model 1: Controls			Model 2: Factors for acceptance		
		b	β	Sig.	b	β	Sig.
Constant		2.129		.000	.412		.155
Control variables							
Age	27.108	-.011	-.055	.136	.001	.004	.894
Gender	.507	-.204	-.065	.053	.014	.004	.863
Education	4.587	.126	.090	.006	.015	.011	.665
Monthly income	2.372	-.008	-.008	.838	-.048	-.043	.122
Residential area	1.667	-.097	-.055	.096	-.058	-.033	.187
Order frequency	4.593	-.110	-.094	.007	-.076	-.064	.014
Amazon Prime customer	0.494	-.033	-.011	.761	-.054	-.017	.509
Delivery concept (standard: reception box)							
Controlled access system	.239	-2.317	-.632	.000	-.085	-.023	.605
Trunk delivery	.248	-1.532	-.423	.000	.026	.007	.844
Crowdsourcing among friends	.251	-1.340	-.372	.000	.170	.047	.259
Factors for acceptance							
Usefulness (H1)	.573				.410	.310	.000
Punctuality (H2)	.709				.037	.028	.486
Flexibility (H3)	1.055				-.004	-.003	.927
Security (H4)	.151				.348	.278	.000
Privacy (H5)	.210				.280	.277	.000
Reliability (H6)	.582				.023	.016	.645
Liability (H7)	.113				.030	.026	.400
Model's summary statistics							
ΔF			31.185			76.533	
Sig. of ΔF			.000			.000	
R^2			.320			.625	

and Cai (2007) to estimate HC-robust standard-errors. The detailed numbers are not reported, as this analysis reproduced the results from the standard OLS estimation regarding the examined hypotheses, i.e., the heteroscedasticity does not bias the hypothesis testing.

In the first model, both education and the consumer's order frequency are found to have a significant effect ($\alpha = .05$) on the response variable, while all other consumer-sided control variables do not show significant effects. All ADC assignment variables have a significant effect on the outcome variable. These results are not surprising, as the descriptive results for the willingness to test an ADC show large variations for the four mentioned ADCs. In the second model, the significant effect of the control variable education disappears. The control variable "order frequency" retains its weak but significant negative influence on the acceptance of the delivery concepts: Frequent shoppers show a slightly higher acceptance of ADCs. A possible explanation might be a low satisfaction level of these consumers with current delivery concepts.

All binary variables for the different delivery concepts are nonsignificant in Model 2. This proves that measuring the general acceptance factors for alternative delivery concepts does not depend on the evaluation of a specific concept. However, not all of these influencing factors show significant effects on the consumer's willingness to test a delivery concept. Usefulness ($b = .410$, $p < .001$), perceived privacy ($b = .280$, $p < .001$), and security ($b = .348$, $p < .001$) are strong and significant predictors of the intention of consumers to test ADCs: The more useful a delivery concept is and the fewer privacy and security issues that are perceived, the higher the acceptance of one of these concepts is. In contrast, punctuality, reliability, and liability have very low positive regression coefficients, whereas flexibility has a very low negative coefficient. These factors seem to not significantly affect the explained variable according to the regression model. Thus, for the target group of young e-tailing consumers in Germany, H1, H4, and H5 cannot be rejected, while H2, H3, H6, and H7 must be rejected.

4 Conclusion

To the best of our knowledge, this paper is the first to examine consumer perceptions of four alternative unattended home delivery concepts proposed by businesses and academia to make the last mile more efficient and eco-friendly. We show that among the four ADCs examined, the reception box is perceived as having the greatest usefulness, flexibility, punctuality, security, privacy, reliability, and lack of liability issues. The greatest skepticism among the respondents was in regard to the punctuality of the crowd delivery through friends approach as well as the security and privacy of controlled access systems. Furthermore, the effects of multiple benefits and risk characteristics on the willingness to test an ADC were investigated using a regression analysis. The model we used achieved a surprisingly high explanatory power, with 63% of the variance being explained. The results show that perceived usefulness, security, and privacy are determining factors for the willingness to test and use ADCs.

With regard to theoretical contributions, this study confirms the results of Inhoffen's (2019) survey. He reported low acceptance levels for controlled access systems. Accordingly, only 22% of the respondents could imagine actually using this delivery concept. In addition, this paper complements the study of Yuen et al. (2018) who investigated the acceptance of self-collection services in the last mile. This article extends this line of research to alternative unattended home delivery options. Such a complementary study is necessary because, according to a recent PWC (2018) study, only less than half of consumers are willing to use self-collection systems. Therefore, it seems unrealistic that such systems are capable of fully meeting the current challenges. In other words, it is important to additionally consider the ADCs examined in this paper. Furthermore, this article pursues a different theoretical perspective. Drawing on the TAM, the extended TAM, the TRA, and the UTAUT, a two-dimensional research model is developed, which can serve as an impetus for other research on the acceptance of logistical concepts. Since the application of a delivery concept requires its acceptance, it seems important to better understand what drives the consumer to adopt certain concepts. In such an effort, this research contributes to the last mile design/configuration literature by providing indications of each concepts' success potential. Therefore, we suggest that the concepts with higher acceptance levels deserve more attention and should be examined as a priority in the future.

Managerially, this study provides decision-makers for e-tailers and logistics service providers with information on which home delivery option they should focus on. Decisions on the design of the delivery chain are often associated with high investments and must therefore be thoroughly prepared. The data from our study suggest that the reception box has the greatest potential among the four concepts. To translate the high acceptance levels into market success, specific political measures are necessary. In Germany, for instance, the vast majority of all homes are double- or multiparty homes (NAI Global 2019). In addition, a large part of the population rents their accommodations. In such circumstances it is currently very difficult or even impossible to obtain a permit for the construction of a reception box. This is reminiscent of the discussion about private charging stations for electric vehicles. Other countries are much more progressive in this respect. In 2012, Switzerland introduced the obligation to equip letterboxes with a stowage compartment at least for small parcels (Schweizerische Bundeskanzlei 2019). This could act as a blueprint for other policy-makers. Furthermore, there are a large number of service providers competing in the last mile parcel delivery market (e.g., DHL, DPD, UPS, Hermes, GLS, TNT in Germany). It is therefore necessary to define an open standard for securely accessing reception boxes, not least because it is impracticable to set up a box for every logistics service provider.

While our study provides the aforementioned theoretical and managerial implications, we also acknowledge that this study has limitations. First, this research has only focused on four ADCs, though others have also been suggested (e.g., drone delivery). Second, the data were only collected in Germany and were focused on young consumers, which limits generalizability. Third, most of the investigated constructs are operationalized with only one item and the item "usefulness" can be criticized for overlapping with others and not being mutually exclusive. Despite these limitations, we are confident that this study already presents interesting and valuable preliminary results that can serve as impetus for future research. In fact, the authors of this study regard the survey as an explorative pretest and will soon initiate a more comprehensive study in Germany. In addition, studies in other regions (e.g., USA, France, Great Britain) would be of interest in order to identify any international differences.

References

Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)

Amazon: Introducing Amazon Key, A New Level of Delivery Convenience for Prime Members (2017). https://press.aboutamazon.com/news-releases/news-release-details/introducing-amazon-key-new-level-delivery-convenience-prime. Accessed 8 Feb 2019

Amazon: Amazon's Air Network Expands to Support the Growth of Prime Fast, Free Shipping for Customers (2018). https://press.aboutamazon.com/news-releases/news-release-details/amazons-air-network-expands-support-growth-prime-fast-free. Accessed 8 Feb 2019

Anderson, S., Allen, J., Browne, M.: Urban logistics – how can it meet policy makers' sustainability objectives? J. Transp. Geogr. **13**(1), 71–81 (2005)

Audi: Audi, DHL and Amazon Deliver Convenience (2015). https://www.audiusa.com/newsroom/news/press-releases/2015/04/audi-dhl-and-amazon-deliver-convenience. Accessed 8 Feb 2019

Capgemini: The Last-Mile Delivery Challenge: Giving Retail and Consumer Product Customers a Superior Delivery Experience without Impacting Profitability (2019). https://www.capgemini.com/wp-content/uploads/2019/01/Report-Digital-%E2%80%93-Last-Mile-Delivery-Challenge1.pdf. Accessed 10 May 2019

Castillo, V.E., Bell, J.E., Rose, W.J., Rodrigues, A.M.: Crowdsourcing last mile delivery: strategic implications and future research directions. J. Bus. Logist. **39**(1), 7–25 (2017)

Cohen, J.: Statistical Power Analysis for the Behavioral Sciences. Taylor and Francis Inc., Hoboken (1988)

Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. **35**(8), 982–1003 (1989)

Devari, A., Nikolaev, A.G., He, Q.: Crowdsourcing the last mile delivery of online orders by exploiting the social networks of retail store customers. Transp. Res. Part E **105**, 105–122 (2017)

DHL: Logistics Trend Radar – Delivering Insight Today. Creating Value Tomorrow! (2016). https://www.logistics.dhl/content/dam/dhl/global/core/documents/pdf/dhl-logistics-trend-radar-2016.pdf. Accessed 14 Feb 2019

Dowling, G., Staelin, R.: A model of perceived risk and intended risk-handling activity. J. Consum. Res. **21**(1), 119–134 (1994)

eMarketer: Retail E-Commerce Sales Worldwide from 2014 to 2021 (in Billion U.S. Dollars) (2019). https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/. Accessed 8 Feb 2019

Faulin, J., Grasman, S.E., Juan, A.A., Hirsch, P.: Sustainable transportation: concepts and current practices. In: Faulin, J., Grasman, S.E., Juan, A.A., Hirsch, P. (eds.) Sustainable Transportation and Smart Logistics. Decision-Making Models and Solutions, pp. 3–23. Elsevier, Cambridge (2019)

Field, A.: Discovering Statistics Using IBM SPSS Statistics. Sage, London (2013)

Gevaers, R., Van de Voorde, E., Vanelslander, T.: Cost modelling and simulation of last-mile characteristics in an innovative B2C supply chain environment with implications on urban areas and cities. Procedia Soc. Behav. Sci. **125**, 398–411 (2014)

Glesjer, H.: A new test for heteroskedasticity. J. Am. Stat. Assoc. **64**(325), 316–323 (1969)

Glöss, M., McGregor, M., Brown, B.: Designing for labour: Uber and the on-demand mobile workforce. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 1632–1643. ACM, New York (2016)

Hayel, Y., Quadri, D., Jiménez, T., Brotcorne, L.: Decentralized optimization of last-mile delivery services with non-cooperative bounded rational customers. Annu. Oper. Res. **239**(2), 451–469 (2016)

Hayes, A.F., Cai, L.: Using heteroskedasticity-consistent standard error estimators in OLS regression: an introduction and software implementation. Behav. Res. Methods **39**(4), 709–722 (2007)

Inhoffen: Paketzustellung: Jeder Fünfte ist offen für In-Home-Lieferung (2019). https://yougov.de/news/2019/01/17/paketzustellung-jeder-funfte-ist-offen-fur-home-li/. Accessed 22 Feb 2019

Iwan, S., Kijewska, K., Lemke, J.: Analysis of Parcel Lockers' efficiency as the last mile delivery solution – the results of the research in Poland. Transp. Res. Procedia **12**, 644–655 (2016)

Kull, T.J., Boyer, K., Calantone, R.: Last-mile supply chain efficiency: an analysis of learning curves in online ordering. Int. J. Oper. Prod. Manag. **27**(4), 409–434 (2007)

Macharis, C., Melo, S.: Introduction – city distribution: challenges for cities and researchers. In: Macharis, C., Melo, S. (eds.) City Distribution and Urban Freight Transportation, pp. 1–9. Edward Elgar Publishing, Northhampton (2011)

Mandziuk, J.: New shades of the vehicle routing problem: emerging problem formulations and computational intelligence solution methods. IEEE Trans. Top. Comput. Intell. **3**(3), 230–244 (2019)

Martins, C., Oliveira, T., Popovic, A.: Understanding the internet banking adoption: a unified theory of acceptance and use of technology and perceived risk application. Int. J. Inf. Manag. **34**(1), 1–13 (2014)

Marujo, L.G., Goes, G.V., D'Agosto, M.A., Ferreira, A.F., Winkenbach, M., Bandeira, R.A.M.: Assessing the sustainability of mobile depots: the case of urban freight distribution in Rio de Janeiro. Transp. Res. Part D **62**, 256–267 (2018)

McKinnon, A.C., Tallam, D.: Unattended delivery to the home: an assessment of the security implications. Int. J. Retail Distrib. Manag. **31**(1), 30–41 (2003)

Mladenow, A., Bauer, C., Strauss, C.: Crowd logistics: the contribution of social crowds in logistics activities. Int. J. Web Inf. Syst. **12**(3), 379–396 (2016)

Moore, G.C., Benbasat, I.: Development of an instrument to measure the perceptions of adopting an information technology innovation. Inf. Syst. Res. **2**(3), 192–222 (1991)

Morganti, E., Seidel, S., Banquart, C., Dablanc, L., Lenz, B.: The impact of e-commerce on final deliveries: alternative parcel delivery services in France and Germany. Transp. Res. Procedia **4**, 178–190 (2014)

Nabot, A., Omar, F.: Comparative study of the impacts of conventional and online retailing on the environment: a last mile perspective. Int. J. Comput. Appl. **138**(3), 6–12 (2016)

NAI Global: Verteilung der Wohnungen in Deutschland nach Wohngebäuden bis 2016 (2019). http://de.statista.com/statistik/daten/studie/762674/umfrage/verteilung-der-wohnungen-in-deutschland-nach-wohngebaeuden/. Accessed 6 Mar 2019

Ozbaygin, G., Karasan, O.E., Savelsbergh, M., Yaman, H.: A branch-and-price algorithm for the vehicle routing problem with roaming delivery locations. Transp. Res. Part B **100**, 115–137 (2017)

Park, H., Park, D., Jeong, I.-J.: an effects analysis of logistics collaboration in last-mile networks for CEP delivery services. Transp. Policy **50**, 115–125 (2016)

Pavlou, P.A.: consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. Int. J. Electron. Commer. **7**(3), 101–134 (2003)

Punakivi, M., Yrjölä, H., Holmström, J.: Solving the last mile issue: reception box or delivery box? Int. J. Phys. Distrib. Logist. **31**(6), 427–439 (2001)

Punel, A., Stathopoulos, A.: Modeling the acceptability of crowdsourced goods deliveries: role of context and experience effects. Transp. Res. Part E Logist. Transp. Rev. **105**, 18–38 (2017)

PWC: Flexibel, schnell, umweltfreundlich – Wege aus dem Paketdilemma (2018). https://www.pwc.de/de/transport-und-logistik/pwc-paketpreisstudie-2018.pdf. Accessed 25 Feb 2019

Ramadan, Z.B., Farah, M.F., Mrad, M.: An adapted TPB approach to consumers' acceptance of service-delivery drones. Technol. Anal. Strat. Manag. **29**(7), 817–828 (2017)

Reyes, D., Savelsbergh, M., Toriello, A.: Vehicle routing with roaming delivery locations. Transp. Res. Part C **80**, 71–91 (2017)

Robarts, S.: Smart Box Enables Trunk Delivery Trial to Smart Cars (2016). https://newatlas.com/smart-ready-to-drop-trial/44568/. Accessed 14 Feb 2019

Rogers, E.M.: Diffusion of Innovations. Free Press of Glencoe, New York (1962)
Schweizerische Bundeskanzlei: SR 783.01 Postverordnung vom, 29 August 2012 (2019). https://www.admin.ch/opc/de/classified-compilation/20112357/index.html#a73. Accessed 6 Mar 2019
Småros, J., Holmström, J., Kämäräinen, V.: New service opportunities in the e-grocery business. Int. J. Logist. Manag. **11**(1), 61–74 (2000)
Statistisches Bundesamt: Bevölkerungsstand (2019a). https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/zensus-geschlecht-staatsangehoerigkeit-2018.html. Accessed 8 May 2019
Statistisches Bundesamt: Share of Internet Users who Ordered and Purchased Products Online within the Past Three Months in Germany in 2018, by Age Group (2019b). https://www.statista.com/statistics/506181/e-commerce-online-shoppers-by-age-group-germany/. Accessed 7 May 2019
Tan, J.M.: Overcoming the Last-Mile Challenge. Nanyang Technological University, Singapore (2016)
Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.: User acceptance of information technology: toward a unified view. MIS Q. **27**(3), 425–478 (2003)
Volvo: Volvo Expands Its In-Car Delivery Service (2016). https://www.drivesweden.net/en/volvo-expands-its-car-delivery-service. Accessed 8 Feb 2019
Walmart: Why the Future Could Mean Delivery Straight Into Your Fridge (2017). https://blog.walmart.com/innovation/20170922/why-the-future-could-mean-delivery-straight-into-your-fridge. Accessed 14 Feb 2019
Wang, X., Zhan, L., Ruan, J., Zhang, J.: How to choose “Last Mile” delivery methods for e-fulfillment. Math. Probl. Eng. **2014**, 1–11 (2014)
Wang, Y., Zhang, D., Liu, Q., Shen, F., Lee, L.H.: Towards enhancing the last-mile delivery: an effective crowd-tasking model with scalable solutions. Transp. Res. Part E **93**, 279–293 (2016)
Xu, M., Ferrand, B., Roberts, M.: The Last mile of e-commerce – unattended delivery from the consumers and eTailers’ persepctives. Int. J. Electron. Mark. Retail. **2**(1), 20–38 (2008)
Yuen, K.F., Wang, X., Ng, L.T.W., Wong, Y.D.: An investigation of customers’ intention to use self-collection services for last-mile delivery. Transp. Policy **66**, 1–8 (2018)

A Green Supply Chain Design Model Considering Lead Times

Benjamin Siller(✉)

Julius-Maximilians-Universität Würzburg, Sanderring 2, 97072 Würzburg, Germany
benjamin.siller@uni-wuerzburg.de

Abstract. Lead times and carbon emissions are important factors for the design of supply chain networks since companies' customers and various other stakeholders are getting more and more sensitive regarding both performance factors. In order to reduce lead times, companies apply different strategies like e.g. using faster logistic modes, locate production facilities and warehouses near customers or keep goods at stock. Nevertheless, these measures can have a high impact on both costs and carbon emissions. This paper provides a multi-layer, multi-product and multi-period supply chain design approach with a carbon cap-and-trade system considering lead times. The bi-objective model aims to minimize delivery lead time and discounted total costs. Furthermore, it is possible to keep units on stock and carbon emissions for all necessary processes to fulfill customer demand are taken into account. A computational study evaluates the solvability of the model and gives insight on the influence of delivery lead time and prices of carbon credits on the configuration of a supply chain.

Keywords: Supply chain design · Lead times · Carbon cap-and-trade · Mixed integer programming

1 Introduction

Looking for new competitive advantages, time can be the source of success for companies. Customer perception of products and services regarding time-related characteristics are likely to be a source of differentiation [1]. Therefore, companies have discovered lead time reduction as a competitive weapon. Nowadays, supply chains often have complex network structures with multiple layers. Facilities are located all over the world and a network of suppliers from different countries is part of the supply network. Due to geographically dispersed production facilities and warehouses, lead times are important for the success of a global manufacturing strategy [2]. Lead time for a specific order is often defined as the time span between the moment the order is placed and the moment when the order is fulfilled. Lead time usually consists of several components like e.g. supplier lead time, delivery lead time, setup time and order transit time [3]. Customers have a tolerance for the amount of time it takes for the order to be fulfilled. Therefore, it is important for a company to align the production and logistics activities in a way that this tolerance is not exceeded [4]. Furthermore, the influence of lead time on both the performance of distribution systems and inventory control policy

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 172–187, 2019.
https://doi.org/10.1007/978-3-030-29821-0_12

is well documented [5]. In the last decades, customer requirements changed. Today, in many cases specialized and customized products with small lot sizes and short delivery lead times are required [6]. In addition, consumers' concerns regarding climate change are growing and customers become more familiar with the concept of carbon footprint [7]. Also, it is well known that carbon emissions affect global warming [8]. Furthermore, companies are facing stakeholders, who more often demand higher levels of sustainability that also affect the supply chain [9]. One way to reduce carbon emissions are cap-and-trade systems. Such systems are generally accepted as an effective, market-based mechanism to reduce carbon emissions by companies. In such a system, companies initially get free emission credits. These emission credits can be traded between companies on specific carbon trading markets. If the allocated amount is not sufficient for the individual production plans, the company can buy more credits on the market or invest in greener technology to reduce carbon emissions [10].

In general, supply chain management can be classified in operational, tactical and strategic tasks. Operational and tactical tasks deal with decisions in a short to medium planning horizon, whereas strategic decisions address a rather long planning horizon. Strategic decisions cannot be easily modified in short-term and/or only at high costs. Environmental performance is strongly affected by long-term environmental supply chain structure planning [11]. Furthermore, replenishment lead times have a crucial impact on the responsiveness of a supply chain, a fact that became more crucial due to extended lead times in globalized value networks [12]. Therefore, it is important to incorporate both aspects, carbon emissions and lead times, in a strategic decision model.

The purpose of this contribution is to combine lead time and cost trade-offs in a supply chain under a carbon emission cap-and-trade system. After the introduction, in Sect. 2, a brief literature review on supply chain design approaches including lead time and cap-and-trade systems is provided. The problem statement, model assumption, and the model are introduced in Sect. 3. In Sect. 4 this model is evaluated with random data and results are discussed. Finally, conclusions are drawn in Sect. 5.

2 Literature Review

In most studies lead times are used to determine safety stocks (see e.g. [13]), to model time gaps (see e.g. [14]), or for operational or tactical management decisions (see e.g. [15]).

One of the first approaches for incorporating lead times in strategic supply chain design models is developed by Arntzen et al. [16]. In their study, they develop a model, which includes manufacturing and transportation lead time in a single-objective function. The model aims at minimizing costs and determines facility locations in a strategic network. Melachrinoudis and Min [17] propose a multi-objective supply chain design model. The model considers total costs, distribution lead time between hybrid plants/warehouses and customers and local incentives for facilities. Sadjady and Davoudpour [18] propose a model for the design of distribution networks. It includes decisions on locations and the size of manufacturing plants and warehouses as well as the selection of transportation modes. The approach aims to minimize total costs of the network and includes transportation lead time, inventory, operations, and fixed costs. Hammami and Frein [19] present a MIP model for the design of a global supply chain. Quoted lead time is incorporated as a constraint and purchasing, manufacturing and transportation lead times are considered. Liu and Papageorgiou [20] use a multi-objective MILP approach for planning a global supply chain in a process industry context. The objectives include total costs, total flow time, consisting of transportation lead time, and lost sales. The trade-off between supply chain costs and lead times is considered in a bi-objective model by Zhang et al. [21]. In the context of dispersed manufacturing in China, the objectives of the model are minimization of both, cost and lead time to determine the optimal facility location. Martí et al. [22] propose a supply chain design model, which includes responsiveness decisions under carbon policies. Therefore, manufacturing and transportation lead times are incorporated. The addressed carbon policies are caps on supply chain carbon footprint, caps on market carbon footprint and carbon taxes. A deterministic optimization model incorporating carbon emissions and lead time constraints is developed by Hammami et al. [23]. Carbon emission tax and emission cap are discussed in this study. Furthermore, purchasing, manufacturing, and transportation lead times are considered. Meisel et al. [24] investigate the facility network design for the production and transportation of goods under customer lead time restrictions. They provide two alternative models to design a multi-layer, multi-product facility network considering purchasing, manufacturing, and transportation lead times. Make to stock and make to order production strategies are included in both models. Hammami et al. [25] propose a model that guarantees quoted lead time, including purchasing, manufacturing, and transportation lead time for each customer order and the replenishment of stocks in different facilities. With the model, the impact of customer order frequency and quoted lead time on supply chain decisions and costs are investigated.

Even though most models considering lead times do not take carbon policies into account, Carbon emissions are a well discussed topic in the supply chain context. For an overview of recent studies about carbon policies, see [26, 27, 28].

Fareeduddin et al. [29] develop several models with strict carbon cap, carbon tax, and carbon cap-and-trade policies to design a closed-loop supply chain network. The models minimize costs and emissions and capture the trade-offs between facility location and logistic modes, as well as between costs and emissions in supply chain operations. A green supply chain network design problem is solved by Shaw et al. [30]. The chance-constrained model includes uncertainties of capacities and demand and carbon cap-and-trade policy. An optimization model for the design and planning of a multi-product, multi-period closed-loop supply chain considering carbon footprint and uncertainties is developed by Mohammed et al. [31]. Furthermore, the model investigates the influence of different carbon policies (strict carbon cap, carbon tax, carbon cap-and-trade, and carbon offset). A two-stage stochastic programming model for designing a green supply chain under carbon cap-and-trade system is proposed by Rezaee et al. [32]. In this approach, uncertain carbon prices and uncertain demand are considered. A mixed-integer linear programming model for a multi-period bio-energy supply chain is proposed by Memari et al. [33]. The model is applied under carbon tax and carbon cap-and-trade policy.

Table 1. Summary of literature review

Contribution	Procurement lead time	Transportation lead time	Manufacturing lead time	Handling lead time	Lead time implementation	Carbon cap and trade	Carbon cap	Carbon tax	Multi-period	Multi-product	Inventory	Logistic modes
[16]		x	x		Objective				x	x	x	x
[17]		x			Objective				x			
[18]		x			Objective					x	x	x
[19]	x	x	x		Lead time cap						x	
[20]		x			Objective				x	x	x	
[21]		x	x		Objective							
[22]		x	x		Inventory costs		x	x			x	x
[23]	x	x	x		Lead time cap		x	x	x		x	
[24]	x	x	x		Lead time cap					x	x	x
[25]	x	x	x		Lead time cap						x	
[29]					–	x	x	x		x		x
[30]					–	x						
[31]					–	x	x	x	x	x	x	x
[32]					–	x				x		x
[33]					–	x		x	x		x	x

Table 1 summarizes the finding of the literature review. Only in [22] and [23] supply chain design models considering both, lead time and carbon emissions are proposed, but none of these studies includes a cap-and-trade policy. Therefore, a bi-objective multi-period, multi-product supply chain design model with considerations of logistic modes, inventories, lead time and cap-and-trade policy is proposed in the following.

3 Modelling Framework and Assumptions

3.1 Problem Description

In addition to taking into account discounted costs, delivery lead time is considered in a separate objective function. Delivery lead time is the time needed to fulfill a customer's order of a product in a period and consists of transportation lead time, procurement lead time, manufacturing lead time, and handling lead time in the warehouses. If the amount of products in stock is equal or higher than the customer order, delivery lead time consists of handling lead time and transportation lead time from a warehouse to the specific customer. When the customer order is higher than the stock level, procurement lead time, manufacturing lead time and transportation lead time between supplier, production facilities and warehouses for the needed amount of parts that is not in stock have to be taken into account as well as handling lead times for processing units in warehouses. Lead time is seen as lead time per unit as e.g. in [16] and [21].

Furthermore, the developed model is based on the following assumptions:

- The proposed model is a discrete deterministic model with finite numbers of potential suppliers, production facilities, warehouses and planning periods.
- There is a fixed number of customer regions.
- The strategic planning horizon covers several planning periods.
- The model covers several products.
- Capacities of suppliers, production facilities and warehouses are restricted.
- Transportation activities can be done with several logistic modes.

For the formulation of the optimization model, the following notations are used:

Sets

S Set of candidate suppliers indexed by s
F Set of candidate locations for production facilities indexed by f
W Set of candidate locations for warehouses indexed by w
C Set of customers indexed by c
T Set of same length time periods in the planning horizon indexed by t
M Set of raw materials indexed by m
P Set of products indexed by p
L Set of logistic modes indexed by l

Parameters

Dem_{ptc} Demand of customer c for product p in period t
FC_s Fixed costs for contracting supplier s

FC_f	Fixed setup costs of production facility f
FC_w	Fixed setup costs of warehouse w
TC_{sfmtl}	Unit transportation costs for raw material m shipped by logistic mode l from supplier s to production facility f in period t
TC_{fwptl}	Unit transportation costs for product p shipped by logistic mode l from production facility f to warehouse w in period t
TC_{wcptl}	Unit transportation costs for product p shipped by logistic mode l from warehouse w to customer c in period t
PC_{smt}	Unit purchasing costs for raw material m at supplier s in period t
PC_{fpt}	Unit production costs for product p at production facility f in period t
HC_{wpt}	Unit inventory costs for product p at warehouse w in period t
LT_{sfmtl}	Unit transportation lead time for raw material m shipped by logistic mode l from supplier s to production facility f in period t
LT_{fwptl}	Unit transportation lead time for product p shipped by logistic mode l from production facility f to warehouse w in period t
LT_{wcptl}	Unit transportation lead time for product p shipped by logistic mode l from warehouse w to customer c in period t
LT_{smt}	Unit purchasing lead time for raw material m at supplier s in period t
LT_{fpt}	Unit production lead time for product p at production facility f in period t
LT_{wpt}	Unit handling lead time for product p at warehouse w in period t
FE_f	Fixed setup emission value of production facility f
FE_w	Fixed setup emission value of warehouse w
TE_{sfmtl}	Unit transportation emission value for raw material m shipped by logistic mode l from supplier s to production facility f in period t
TE_{fwptl}	Unit transportation emission value for product p shipped by logistic mode l from production facility f to warehouse w in period t
TE_{wcptl}	Unit transportation emission value for product p shipped by logistic mode l from warehouse w to customer c in period t
PE_{fpt}	Unit production emission value for product p at production facility f in period t
HE_{wpt}	Unit handling emission value for product p at warehouse w in period t
Cap_{sm}	Capacity of raw material m at supplier s
Cap_f	Capacity at production facility f
Cap_w	Capacity of warehouse w
BOM_{mp}	Bill of Materials relating materials m to products p
$ECap_t$	Emission cap in period t
$WACC$	Weight average cost of capital
$ECost$	Price of one carbon credit

Variables

Y_s	1 if supplier s is selected, 0 otherwise
Y_f	1 if a production facility is located at candidate location f, 0 otherwise
Y_w	1 if a warehouse is located at candidate location w, 0 otherwise
X_{sfmtl}	Amount of raw material m shipped from supplier s to production facility f in period t with logistic mode l

X_{fwptl}	Amount of product p shipped from production facility f to warehouse w in period t with logistic mode l
X_{wptlc}	Amount of product p shipped from warehouse w to customer c in period t with logistic mode l
X_{fpt}	Amount of manufactured products p in facility f in period t
h_{wpt}	Amount of product p kept in stock at warehouse w in period t
d_{sfmtlc}	Amount of raw material m from supplier s transported with logistic mode l that is needed at facility f to produce consumer c's order in period t
d_{fwptlc}	Amount of product p, shipped by logistic mode l form production facility f to warehouse w to fulfill the order of customer c in period t
$TrEM_t^-$	Amount of carbon credits for sale in period t
$TrEM_t^+$	Amount of additional carbon credits needed in period t
DLT_{ptc}	Delivery lead time for product p to customer c in period t

3.2 Mathematical Formulation

$$\begin{aligned} Z_1 = \sum\nolimits_t^T \frac{1}{(1+WACC)^t}\Big(&\sum\nolimits_s^S FC_sY_s + \sum\nolimits_f^F FC_fY_f + \sum\nolimits_w^W FC_wY_w + \sum\nolimits_s^S\sum\nolimits_f^F\sum\nolimits_m^M\sum\nolimits_l^L X_{sfmtl}\left(TC_{sfmtl}+PC_{smt}\right) \\ &+ \sum\nolimits_f^F\sum\nolimits_w^W\sum\nolimits_p^P\sum\nolimits_l^L X_{fwptl}TC_{fwptl} + \sum\nolimits_w^W\sum\nolimits_p^P\sum\nolimits_l^L\sum\nolimits_c^C X_{wptlc}TC_{wcptl} \\ &+ \sum\nolimits_f^F\sum\nolimits_p^P X_{fpt}PC_{fpt} + \sum\nolimits_w^W\sum\nolimits_p^P \frac{\left(h_{wpt-1}+h_{wpt}\right)}{2}HC_{wpt} + \left(TrEM_t^+ - TrEM_t^-\right)ECost \end{aligned} \tag{1}$$

$$Z_2 = \sum\nolimits_p^P\sum\nolimits_t^T\sum\nolimits_c^C DLT_{ptc} \tag{2}$$

Objective function (1) minimizes the financial (cost) objective given by the sum of the discounted total costs regarding supply chain decisions for each period. It includes fix costs for selecting suppliers, production facilities and warehouses, supply, production, inventory and distribution costs as well as the costs or yields of buying or selling additional carbon credits. The second objective function (2) minimizes the sum of the delivery lead times for each product to each customer in each period. It is therefore assumed, that there are no customer-related prioritizations of the respective lead times.

$$\begin{gathered} Dem_{ptc} = \sum\nolimits_w^W\sum\nolimits_l^L X_{wptlc} \\ \forall p \in P, t \in T, c \in C \end{gathered} \tag{3}$$

$$\begin{gathered} \sum\nolimits_l^L X_{wptlc} + h_{wpt} = \sum\nolimits_f^F\sum\nolimits_l^L X_{fwptl} + h_{wpt-1} \\ \forall w \in W, p \in P, t \geq 1 \in T \end{gathered} \tag{4}$$

$$\sum\nolimits_{w}^{W}\sum\nolimits_{l}^{L}X_{fwptl}=X_{fpt} \quad \forall f\in F,p\in P,t\in T \tag{5}$$

$$\sum\nolimits_{p}^{P}X_{fpt}BOM_{mp}=\sum\nolimits_{s}^{S}\sum\nolimits_{l}^{L}X_{sfmtl} \quad \forall f\in F,m\in M,t\in T \tag{6}$$

$$\sum\nolimits_{f}^{F}\sum\nolimits_{l}^{L}X_{sfmtl}\le Cap_{sm}Y_{s} \quad \forall s\in S,m\in M,t\in T \tag{7}$$

$$\sum\nolimits_{p}^{P}X_{fpt}\le Cap_{f}Y_{f} \quad \forall f\in F,t\in T \tag{8}$$

$$\sum\nolimits_{p}^{P}h_{wpt}\le Cap_{w}Y_{w} \quad \forall w\in W,t\in T \tag{9}$$

$$\sum\nolimits_{p}^{P}\sum\nolimits_{l}^{L}X_{wptl}\le Cap_{w}Y_{w} \quad \forall w\in W,t\in T \tag{10}$$

$$\sum\nolimits_{l}^{L}\sum\nolimits_{c}^{C}X_{wptlc}-h_{wpt-1}\le\sum\nolimits_{f}^{F}\sum\nolimits_{l}^{L}\sum\nolimits_{c}^{C}d_{fwptlc} \quad \forall w\in W,p\in P,t\ge 1\in T \tag{11}$$

$$X_{fwptl}\ge\sum\nolimits_{c}^{C}d_{fwptlc} \quad \forall f\in F,w\in W,p\in P,t\in T,l\in L \tag{12}$$

$$\sum\nolimits_{s}^{S}\sum\nolimits_{m}^{M}\sum\nolimits_{l}^{L}\sum\nolimits_{c}^{C}d_{sfmtlc}BOM_{mp}=\sum\nolimits_{w}^{W}\sum\nolimits_{l}^{L}\sum\nolimits_{c}^{C}d_{fwptlc} \quad \forall f\in F,m\in M,t\in T \tag{13}$$

$$X_{sfmtl}\ge\sum\nolimits_{c}^{C}d_{sfmtlc} \quad \forall s\in S,f\in F,p\in P,t\in T \tag{14}$$

$$\begin{aligned}DLT_{ptc}\ge{}& LT_{sfmtl}d_{sfmtlc}+LT_{fwptl}d_{fwptlc}+LT_{wcptl}X_{wptlc}+LT_{smt}d_{sfmtlc}\\&+LT_{fpt}d_{fwptlc}+LT_{wt}X_{wtc}\\&\forall s\in S,f\in F,w\in W,m\in M,p\in P,t\in T,l\in L,c\in C\end{aligned} \tag{15}$$

$$\begin{aligned}
&\sum\nolimits_{f}^{F} FE_f Y_f + \sum\nolimits_{w}^{W} FE_w Y_w + \sum\nolimits_{s}^{S}\sum\nolimits_{f}^{F}\sum\nolimits_{m}^{M}\sum\nolimits_{l}^{L} TE_{sfmtl}X_{sfmtl} \\
&\quad + \sum\nolimits_{f}^{F}\sum\nolimits_{w}^{W}\sum\nolimits_{p}^{P}\sum\nolimits_{l}^{L} TE_{fwptl}X_{fwptl} + \sum\nolimits_{w}^{W}\sum\nolimits_{p}^{P}\sum\nolimits_{l}^{L}\sum\nolimits_{c}^{C} TE_{wptlc}X_{wptlc} \\
&\quad + \sum\nolimits_{f}^{F}\sum\nolimits_{p}^{P} PE_{fpt}X_{fpt} + \sum\nolimits_{w}^{W}\sum\nolimits_{p}^{P} HE_{wpt}h_{wpt} + TrEM_t^{-} = \mathrm{ECap}_t + TrEM_t^{+} \\
&\quad \forall\, t \in T
\end{aligned} \tag{16}$$

$$\begin{gathered}
X_{sfmtl}, X_{fwptl}, X_{wptlc}, X_{fpt}, h_{wpt}, d_{sfmtlc}, d_{fwptlc}, TrEM_t^{-}, TrEM_t^{+}, DLT_{ptc} \\
\forall\, s \in S, f \in F, w \in W, m \in M, p \in P, t \in T, l \in L, c \in C
\end{gathered} \tag{17}$$

$$\begin{gathered}
Y_s, Y_f, Y_w \in \{0, 1\} \\
\forall\, s \in S, f \in F, w \in W
\end{gathered} \tag{18}$$

Constraint (3) ensures that customer demand is fulfilled in every period. In (4) flows between production facilities and warehouses are determined as well as the inventory stocks at each warehouse. Constraint (5) determines the production amount and (6) ensures that material flows from suppliers are calculated according to the relevant bill of materials. Constraints (7)–(10) are capacity constraints for suppliers, production facilities, and warehouses. In (11) the number of additional goods, which are not on stock and are needed to fulfill specific customer demand in period t, is calculated. Constraint (12) ensures that the additional products are shipped within the product flows between production facilities and warehouses. Constraint (13) calculates the needed additional raw materials from suppliers to fulfill customer demand in t and (14) makes certain that these raw materials are within the flow from the suppliers to production facilities. In (15) maximum delivery lead times for products p in periods t for customers c are determined. Constraint (16) determines the number of carbon credits to sell or buy. The difference between the sum of fixed emissions of production facilities and warehouses and variable emissions for distribution, production and warehousing, and the emission cap are the number of carbon credits, which have to be bought or sold. Constraint (17) ensures that variables are non-negative and (18) determines the binary variables.

Common used methods of multi-objective optimization are the techniques of Goal Programming, Compromise Programming and the Reference Point Method [34]. To combine the two objective functions, Compromise Programming is applied. This technique offers the advantage that the decision maker has only to know about the relative preferences of the objective weights and is often used in multi-objective approaches. Objective (1) and (2) do not have the same scale, so for the proper use of compromise programming, their values have to be normalized [35]. Based on the Lp-metrics method, it is assumed that the optimal value of Z_1 is Z_1^* and the optimal value of Z_2 is Z_2^*. The Lp-metrics objective function can be formulated as follow:

$$Min\ Z_3 = \left[\varpi \frac{Z_1 - Z_1^*}{Z_1^*} + (1 - \varpi)\frac{Z_2 - Z_2^*}{Z_2^*}\right]. \tag{19}$$

The relative weight of the objective components is given by $0 \leq \varpi \leq 1$ and ϖ is chosen by the decision maker. Using (19) as objective and constraints (3) to (18), a mixed integer linear programming model with a single objective is created.

4 Numerical Examples and Results

4.1 Data Generation

In order to evaluate the model, it is applied to random data sets. These data sets are generated as follows.

For the location of suppliers, production facilities, warehouses and customers, an approach proposed by Melkote and Daskin [36] is used. A node on a 100 × 100 grid represents every location. The x and y coordinates are randomly generated within a uniform distribution U[0, 100] and distance is measured with Euclidean metric. Transportation costs for materials are calculated by dividing the distance by a random number of uniform distribution U[85, 100], for products by dividing by a random number of uniform distribution U[75, 90]. Transport emissions are calculated by distance divided by a random number from U[70, 90] for materials and U[65, 95] for products. Transportation lead time for materials is calculated by dividing the distance by a random integer from range [6, 9], for products by a random integer from range [5, 8]. The results are multiplied with a distinct factor for logistic modes. Demand is generated by using a random integer from range [50, 400] for each product, period and customer. For the Bill of Material, random integer numbers from range [0, 3] are assigned to each product and raw material. At the end of this process, it is checked that every material is assigned to at least one product and that each product needs at least one raw material. Capacity for warehouses and production facilities is generated by dividing the maximum demand by the number of warehouses/production facilities. The result is multiplied by a random number from uniform distribution U[1, 3]. Every material is assigned to a supplier with a 50% chance. Capacity for suppliers is determined by dividing the needed raw material by the number of suppliers providing this raw material. The capacity is the sum over these values for every raw material a supplier is providing multiplied by a random number from uniform distribution U[1, 2].

Most fixed costs are dependent on the capacity of a facility. Therefore a formula similar to the suggestion by Cortinhal and Captivo [37] is used:

$$f_i = U[x_1, y_1] + U[x_2, y_2] * \sqrt{a_i}$$

$U[x_1, y_1]$ and $U[x_2, y_2]$ are uniform distributions in the range $[x_1, y_1]$ and $[x_2, y_2]$ and a_i is the capacity of the specific facility or resource. For the data in this paper, the following formulations are used:

- Fixed costs for contracting a supplier: $U[1000, 2000] + U[10, 20] * \sqrt{a_i}$
- Fixed setup costs of a production facility: $U[10000, 20000] + U[85, 140] * \sqrt{a_i}$
- Fixed setup costs of a warehouse: $U[10000, 20000] + U[10, 25] * \sqrt{a_i}$
- Fixed emission value of a production facility: $U[4000, 8000] + U[30, 75] * \sqrt{a_i}$
- Fixed emissions of a warehouse: $U[3000, 6000] + U[20, 60] * \sqrt{a_i}$

For determining the emission cap of period 1, the mean value for all activities multiplied by maximum demand and the sum of emissions from production facilities and warehouses is reduced by 20%. For every further period the emission cap is reduced by an additional 5%.

Distributions for procurement, production and inventory/handling costs and emissions can be derived from Table 2.

Table 2. Data generation of cost and emission

	Procurement	Production	Inventory/Handling
Cost	U[3,6]	U[4, 8]	U[2, 4]
Emission		U[2, 5]	U[1, 3]

Procurement, production, and handling lead times are assumed to be cost dependent. Therefore, the lead times are calculated by dividing the actual cost value by the maximum cost value. To depict the period dependent fluctuations, all variable parameters are multiplied with a random number of uniform distribution U[0.9, 1.1] for each period. The stocks for period 0 are determined by dividing the maximum total demand for one product by the number of warehouses and multiplying it by a random number of uniform distribution U[0.15, 0.45].

4.2 Results

The test problem is solved by IBM ILOG CPLEX Optimizer 12.80 on an Intel Core i7 - 2.2 GHz with 16 GB Ram. The test problem includes 6 supplier, 4 potential locations for production facilities, 4 potential locations for warehouses and 16 customer regions. Furthermore, 4 raw materials, 2 products, 3 periods and WACC is set to 3%. The price of carbon credits is set to 15.

Figure 1 illustrates the results of this problem and shows the effect of changing the weight parameter of the objective functions. As can be seen in Fig. 1 with an increase of the weight parameter ϖ costs decrease. For $\varpi = 0.9$, costs are nearly 60% less compared to the value for $\varpi = 0.1$. Delivery lead time increases with an increasing weight factor. The highest increase is between $\varpi = 0.8$ and $\varpi = 0.9$. This fact can be explained by keeping more units on stock, using more expensive logistic modes, contracting more suppliers and opening more production facilities and warehouses within scenarios with lower weighting factor.

In addition, emissions depend on the weighting factor, as shown in Fig. 2. In this Figure, it can be observed that with an increasing weighting factor, the number of emission credits decreases significantly. For $\varpi = 0.6$ emission credits can be sold in period 1 and 2. In the case of $\varpi = 0.9$, only 2527 emission credits must be bought in period 3, whereas in period 1 and 2 emission credits can be sold. Therefore, with higher weighting factors the emission trading system can generate an additional source of income. Units on stock are a relevant explanation for this effect. With a low weighting

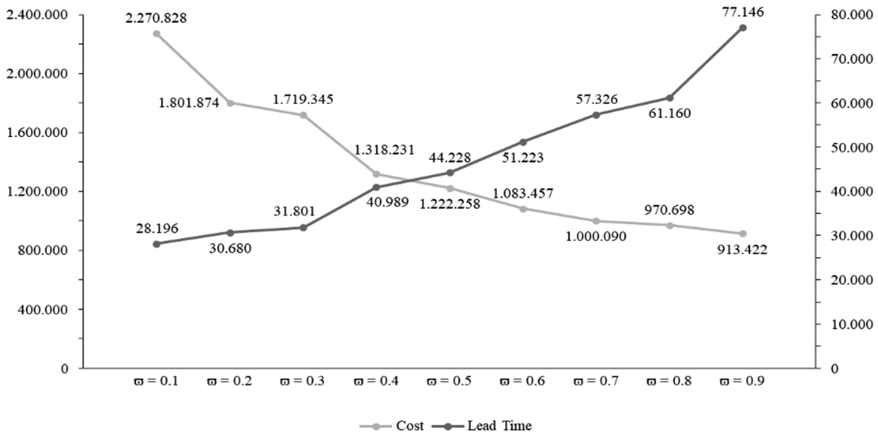


Fig. 1. Comparison of costs and delivery lead time with varying weighting factors.

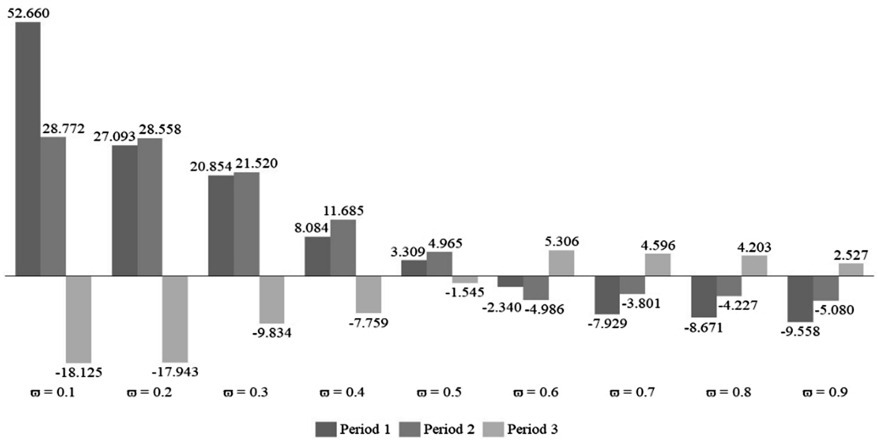


Fig. 2. Comparison of additionally needed and sold emission credits with varying weighting factor.

factor in early periods much more products are produced to raise stocks and lower lead-times in future periods.

Another important parameter for the model is the price of emission credits. Figure 3 shows the changes in cost and delivery lead time with changing emission credit price. The weighting factor is set to $\varpi = 0.5$.

With increasing emission credit price delivery lead time increases. This is due to producing fewer products on stock for future periods, using less emission-intensive logistic modes and opening fewer production facilities and warehouses. Delivery lead times increase between carbon emission price 0 and 30 by more than 115%. Therefore, considering emissions in the model has a high effect on designing the supply chain.

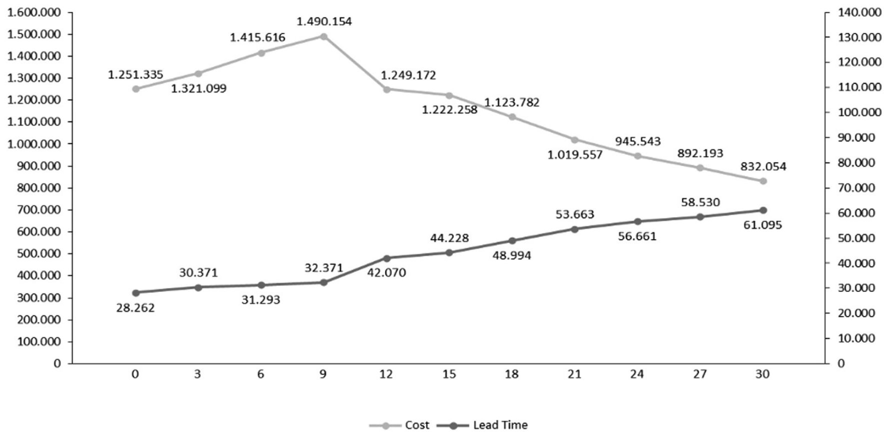


Fig. 3. Comparison of costs and delivery lead time with varying emission credit price.

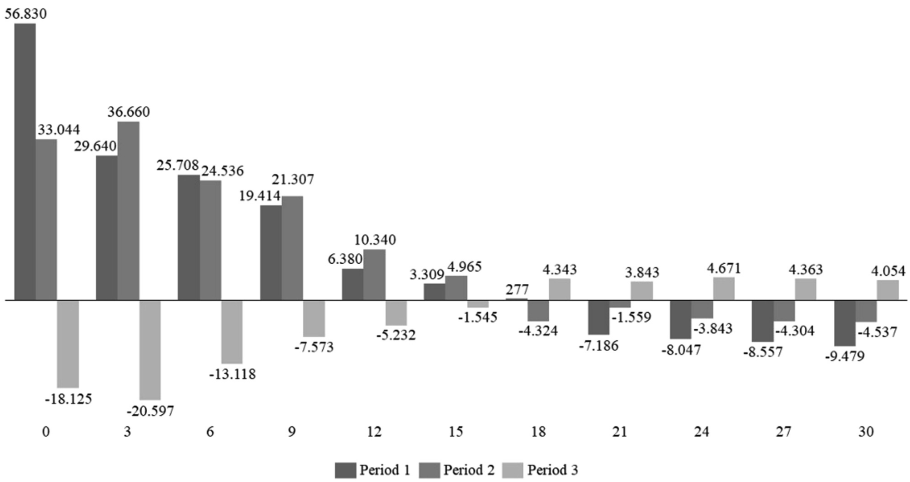


Fig. 4. Comparison of additionally needed and sold emission credits with varying emission credit price.

Costs increase between a carbon emission credit price of 0 and 9. With further increasing credit price costs start to decrease. Figure 4 shows the number of emission credits, which have to be bought or could be sold. With an emission credit price of 21 more emission credits can be sold that have to be bought within all periods. This is an additional source of income and lowers cost. Without considering lead times even more emission credits could be sold and higher emission credit prices lower the cost value of Z_1^* additionally. Therefore, the gap between the cost value (Z_1) and optimal cost value (Z_1^*) increases with higher emission credit price. With overall optimal lead times (Z_2^*) being not directly affected by the emission credit price also the gap between, with higher emission credit price increasing, actual lead time (Z_1) and overall optimal lead time (Z_2^*) is also growing. This fact can explain the increasing overall objective value.

Although the results depend highly on the planning data, it can be stated that the weighting factor and the emission credit price have a high impact on the configuration of the supply chain.

5 Conclusion

Delivery lead times are often seen as competitive weapons and customers more and more expect to receive their orders as fast as possible. On the other side, carbon emissions play an important role in the configuration of a supply chain. Due to governmental regulations like carbon cap, carbon tax, or carbon cap-and-trade, nowadays, carbon emissions must be considered in the designing phase of a supply chain. Carbon cap-and-trade systems are widely accepted as the most effective system to control emissions. To face these problems, a multi-period, multi-product supply chain design model, which minimizes both, total discounted costs and delivery lead times, has been developed. The two objectives are usually conflicting. Furthermore, a carbon cap-and-trade system is implemented to cover environmental considerations. To the authors best knowledge, there is no approach yet that considers order lead time as an objective function and that models a cap-and-trade system. The model is evaluated with randomly generated datasets. The results of the problem illustrate the impact of the discussed factors.

Nevertheless, this study has some limitations. No real-world data has been applied and the model underlies some simplifications. Therefore, only preliminary results could be presented here. Further developments might include several industry cases, to test the approach with real-world data and to verify the computed results. Furthermore, taking into account global economic dynamics, like taxes, tariffs or local content regulations, might be a worthwhile extension. In addition, taking uncertainties, e.g. demand or lead time uncertainties, into account could be an interesting field to extend the presented model.

References

1. de Toni, A., Meneghetti, A.: Traditional and innovative paths towards time-based competition. Int. J. Prod. Econ. **66**(3), 255–268 (2000)
2. Fawcett, S.E.: Strategic logistics in co-ordinated global manufacturing success. Int. J. Prod. Res. **30**(5), 1081–1099 (2007)
3. Tersine, R.J.: Principles of Inventory and Materials Management, 4th edn. Prentice Hall, Englewood Cliffs (1994)
4. Holweg, M.: The three dimensions of responsiveness. Int. J. Oper. Prod. Manag. **25**(7), 603–622 (2005)
5. Berling, P., Farvid, M.: Lead-time investigation and estimation in divergent supply chains. Int. J. Prod. Econ. **157**, 177–189 (2014)

6. Bianchini, A., Benci, A., Pellegrini, M., Rossi, J.: Supply chain redesign for lead-time reduction through Kraljic purchasing portfolio and AHP integration. Benchmarking: Int. J. **26**(4), 1194–1209 (2019)
7. Groening, C., Inman, J.J., Ross Jr., W.T.: The role of carbon emissions in consumer purchase decisions. Int. J. Environ. Policy Decis. Mak. **1**(4), 261–296 (2015)
8. Xu, X., Xu, X., He, P.: Joint production and pricing decisions for multiple products with cap-and-trade and carbon tax regulations. J. Clean. Prod. **112**, 4093–4106 (2016)
9. Meixell, M.J., Luoma, P.: Stakeholder pressure in sustainable supply chain management. Int. J. Phys. Distrib. Logist. Manag. **45**(1/2), 69–89 (2015)
10. Xu, X., He, P., Xu, H., Zhang, Q.: Supply chain coordination with green technology under cap-and-trade regulation. Int. J. Prod. Econ. **183**, 433–442 (2017)
11. Harris, I., Naim, M., Palmer, A., Potter, A., Mumford, C.: Assessing the impact of cost optimization based on infrastructure modelling on CO2 emissions. Int. J. Prod. Econ. **131**(1), 313–321 (2011)
12. Christopher, M., Peck, H., Towill, D.: A taxonomy for selecting global supply chain strategies. Int. J. Logis. Manag. **17**(2), 277–287 (2006)
13. Sourirajan, K., Ozsen, L., Uzsoy, R.: A genetic algorithm for a single product network design model with lead time and safety stock considerations. Eur. J. Oper. Res. **197**(2), 599–608 (2009)
14. Robinson, A.G., Bookbinder, J.H.: NAFTA supply chains. Facilities location and logistics. Int. Trans. Oper. Res. **14**(2), 179–199 (2007)
15. Kaminsky, P., Kaya, O.: Inventory positioning, scheduling and lead-time quotation in supply chains. Int. J. Prod. Econ. **114**(1), 276–293 (2008)
16. Arntzen, B.C., Brown, G.G., Harrison, T.P., Trafton, L.L.: Global supply chain management at digital equipment corporation. Interfaces **25**(1), 69–93 (1995)
17. Melachrinoudis, E., Min, H.: The dynamic relocation and phase-out of a hybrid, two-echelon plant/warehousing facility. A multiple objective approach. Eur. J. Oper. Res. **123**(1), 1–15 (2000)
18. Sadjady, H., Davoudpour, H.: Two-echelon, multi-commodity supply chain network design with mode selection, lead-times and inventory costs. Comput. Oper. Res. **39**(7), 1345–1354 (2012)
19. Hammami, R., Frein, Y.: An optimisation model for the design of global multi-echelon supply chains under lead time constraints. Int. J. Prod. Res. **51**(9), 2760–2775 (2013)
20. Liu, S., Papageorgiou, L.G.: Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. Omega **41**(2), 369–382 (2013)
21. Zhang, A., Luo, H., Huang, G.Q.: A bi-objective model for supply chain design of dispersed manufacturing in China. Int. J. Prod. Econ. **146**(1), 48–58 (2013)
22. Martí, J.M.C., Tancrez, J.-S., Seifert, R.W.: Carbon footprint and responsiveness trade-offs in supply chain network design. Int. J. Prod. Econ. **166**, 129–142 (2015)
23. Hammami, R., Nouira, I., Frein, Y.: Carbon emissions in a multi-echelon production-inventory model with lead time constraints. Int. J. Prod. Econ. **164**, 292–307 (2015)
24. Meisel, F., Rei, W., Gendreau, M., Bierwirth, C.: Designing supply networks under maximum customer order lead times. IIE Trans. **48**(10), 921–937 (2015)
25. Hammami, R., Frein, Y., Bahli, B.: Supply chain design to guarantee quoted lead time and inventory replenishment. Model and insights. Int. J. Prod. Res. **55**(12), 3431–3450 (2016)
26. Das, C., Jharkharia, S.: Low carbon supply chain. A state-of-the-art literature review. J. Manuf. Technol. Manag. **35**(4), 313 (2018)
27. Memari, A., Rahim, A.R.A., Ahmad, R., Hassan, A.: A literature review on green supply chain modelling for optimising CO2 emission. Int. J. Oper. Res. **26**(4), 509 (2016)

28. Waltho, C., Elhedhli, S., Gzara, F.: Green supply chain network design. A review focused on policy adoption and emission quantification. Int. J. Prod. Econ. **208**, 305–318 (2019)
29. Fareeduddin, M., Hassan, A., Syed, M.N., Selim, S.Z.: The impact of carbon policies on closed-loop supply chain network design. Procedia CIRP **26**, 335–340 (2015)
30. Shaw, K., Irfan, M., Shankar, R., Yadav, S.S.: Low carbon chance constrained supply chain network design problem, A Benders decomposition based approach. Comput. Ind. Eng. **98**, 483–497 (2016)
31. Mohammed, F., Selim, S.Z., Hassan, A., Syed, M.N.: Multi-period planning of closed-loop supply chain with carbon policies under uncertainty. Transp. Res. Part D: Transp. Environ. **51**, 146–172 (2017)
32. Rezaee, A., Dehghanian, F., Fahimnia, B., Beamon, B.: Green supply chain network design with stochastic demand and carbon price. Ann. Oper. Res. **250**(2), 463–485 (2017)
33. Memari, A., Ahmad, R., Abdul Rahim, A.R., Akbari Jokar, M.R.: An optimization study of a palm oil-based regional bio-energy supply chain under carbon pricing and trading policies. Clean Technol. Environ. Policy **20**(1), 113–125 (2018)
34. Romero, C., Tamiz, M., Jones, D.F.: Goal programming, compromise programming and reference point method formulations. Linkages and utility interpretations. J. Oper. Res. Soc. **49**(9), 986–991 (2017)
35. Ravindran, A. (ed.): Operations research methodologies. The operations research series. CRC Press, Boca Raton (2009)
36. Melkote, S., Daskin, M.S.: Capacitated facility location/network design problems. Eur. J. Oper. Res. **129**(3), 481–495 (2001)
37. Cortinhal, M.J., Captivo, M.E.: Upper and lower bounds for the single source capacitated location problem. Eur. J. Oper. Res. **151**(2), 333–351 (2003)

Emission Oriented vs. Time Oriented Routing in the European Intermodal Rail/Road Freight Transportation Network

Arne Heinold(✉) and Frank Meisel

School of Economics and Business, Kiel University, Kiel, Germany
{arne.heinold,meisel}@bwl.uni-kiel.de

Abstract. This study compares emissions and transit times from an environmentally oriented and a time oriented routing of large freight shipments in the European rail/road transportation network. We use the terminal-and-service selection problem (TSSP) to find the optimal routings under the different objectives. We show that substantial differences exist between the emission oriented routing and the time oriented routing. A large-scale simulation study reveals that shipments in the emission minimizing routing emit on average almost half as much emissions as if they were routed with the objective to minimize transit time. At the same time, the average transit time of shipments in the emission oriented routing almost triples compared to the transit time in the time optimal routing. This shows by experiment that substantial emission reductions can be achieved in the European freight transport sector by a corresponding routing of shipments but that this comes at the cost of a much lower service quality.

Keywords: Emission rates · Transit time · European rail/road network · Intermodal transportation · Mesoscopic model

1 Introduction

Recent studies show that intermodal rail/road freight transportation is usually less harmful to the environment than unimodal road-only transportation, see for example de Miranda Pinto et al. (2018) or Heinold and Meisel (2018). However, using intermodal rail/road transportation also impacts performance measures like transit time or cost. Quantifying the trade-off between these measures is required in order to make informed transportation decisions, but literature provides little in this regard so far. In particular, there is a need for more analysis of large-scale transport systems. In this study, we compare results from an environmentally oriented routing and a time oriented routing within the rail corridors from the Trans-European Transport Network (TEN-T). Our results are based

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 188–202, 2019.
https://doi.org/10.1007/978-3-030-29821-0_13

on fixed-sized shipments with random origins and destinations across continental Europe. For these orders, we design a terminal-and-service selection problem (TSSP) to find a desirable path through the considered rail/road network. In total, our network covers more than 11,400 rail services per week and connects more than 19 European countries. The results of our simulation show that there are substantial differences between the emission oriented and the time oriented routings. These differences can be observed for absolute measures of emissions and transit time as well as for distance adjusted rates per kilometer.

The paper is organized as follows. Section 2 briefly discusses relevant literature and Sect. 3 describes the considered problem. In Sect. 4, we present the mathematical notations, the optimization model and the used network data. Section 5 describes the results from our simulation experiments. Section 6 concludes this paper.

2 Literature Review

In this section, we provide a short overview of selected studies that focus on long-distance road, rail and/or intermodal rail/road transportation with objectives that are either related to the environment and/or related to time. Due to this scope of our study, we do not discuss here related paper that address different transport settings, such as emission minimization in urban transportation, see Ehmke et al. (2016). Table 1 shows an overview of those studies that we discuss in more detail.

Table 1. Selected relevant literature for this study.

References	Type	Size[a]	Modes		Objectives	
			Road	Rail	Envir.	Time
Figliozzi (2010)	EVRP	med.	√	-	√	√
Bektaş and Laporte (2011)	PRP	small	√	-	√	√
Bauer et al. (2010)	SND	small	√	√	√	√
Demir et al. (2016)	GISND-TTU	med.	√	√	√	√
Lam and Gu (2016)	INO	med.	√	√	-	√
Winebrake et al. (2008)	SPP	small	√	√	√	√
Heinold and Meisel (2018)	SPP	large	√	√	√	-
de Miranda Pinto et al. (2018)	SPP	small	√	√	√	-
Kim and van Wee (2014)	SPP	small	√	√	√	-
This study	TSSP	large	√	√	√	√

[a]instance size, such as number of nodes or number of rail services

Figliozzi (2010) formulates the emissions vehicle routing problem (EVRP) with time windows, capacity constraints and time-dependent travel times. The EVRP considers emissions either as part of a multi-objective function or as a secondary objective in a hierarchical approach. Bektaş and Laporte (2011) present the pollution routing problem (PRP), which is based on a vehicle routing problem with time windows for homogeneous trucks. The authors consider several objectives, such as minimizing distance, travel time or greenhouse gas emissions. The mentioned studies consider merely the road mode. For a detailed review of green vehicle routing problems we refer to Lin et al. (2014). Bauer et al. (2010) present a service network design (SND) formulation that considers greenhouse gas emissions as a primary objective with both modes, road and rail. The authors test their model on a small real-world network between Austria and Poland. Demir et al. (2016) also investigate a service network design problem and introduce the green intermodal service network design problem with travel time uncertainty (GISND-TTU). The authors apply their model on a small intermodal network, covering trains, trucks and inland vessels, and compare results for the objectives time and emissions. Lam and Gu (2016) present a bi-objective intermodal network optimization (INO) for hinterland transportation, including the road and the rail mode. The authors use cost and time as objectives and model governmental carbon emission restrictions per transport unit as additional constraints in their mathematical model. Winebrake et al. (2008) describe the geospatial intermodal freight transport (GIFT) model, which allows a path evaluation for the dimensions time and emissions. The model is demonstrated as a shortest path problem (SPP) in three case studies in northern USA. A recent study of Heinold and Meisel (2018) uses a simulation approach to estimate emissions of the shortest path for road-only and intermodal rail/road transportation for 27 countries in Europe. The authors show that intermodal transportation is less harmful to the environment for more than 90% of the simulated orders. However, the selection of the used transshipment terminals at the origin area and at the destination area is solely based on distance, which might not be optimal if scheduled rail services have to be considered and transit time is to be minimized. Similarly, de Miranda Pinto et al. (2018) use the nearest train station in a shortest path based case study for the Brazilian paper and cellulose pulp sector. The authors show that intermodal transportation could reduce emissions by up to 77.4% compared to road-only transportation. Kim and van Wee (2014) assess shortest paths with emissions in road-only, intermodal rail/road and intermodal vessel/road transports between Rotterdam and Gdansk. The authors use fix road-distances of at most 50 km for road transportation at each end of the intermodal transport chain. In contrast, the TSSP in our study considers multiple transshipment terminals for the orders and, thus, includes a terminal selection problem when searching for the time- and emission-minimizing routings of freight shipments all across Europe.

3 Problem Description

The problem addressed in this study is to find optimal paths for orders with origins and destinations across Europe. For an individual order, the problem corresponds to a shortest path problem in a directed network with arc weights corresponding to transit times and/or emissions. Thereby multiple time-sensitive arcs can exist between two nodes as there might be multiple rail services offered

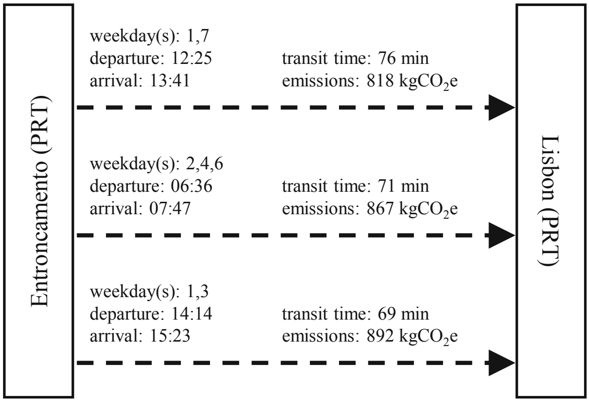


Fig. 1. Representation of rail services as multiple time-sensitive arcs.

between two terminals, see Fig. 1. In the figure, the three arcs represent three scheduled rail services between the two Portuguese rail stations Entroncamento and Lisbon. The trains operate on different weekdays, depart and arrive at different times and emit different amounts of greenhouse gases. The latter is because transit time is closely related to speed, which is an important factor in estimating emissions from rail transportation.

Common shortest path algorithms, such as the algorithms of Dijkstra (1959) or Floyd (1962), cannot be used to solve this kind of problem because they cannot deal with multiple time-dependent arcs between a pair of nodes. Instead, we model the problem addressed in this paper as a special case of a service network design problem, see Crainic (2000). Here, we consider only one order at a time and assume enough capacity for this order for all services in the rail network. We name this modified service network design problem a terminal-and-service selection problem (TSSP). Clearly, the routing of an individual order could also be considered as a time-dependent shortest path problem. Anyhow, we model it here as a variant of the service network design problem because the TSSP can be extended towards multiple orders in future research easily. For the considered order, we measure emissions for the path that minimizes total transit time and measure the transit time for the path that minimizes total emissions. In other words, we solve the TSSP once under the objective to minimize transit time and once under the objective to minimize emissions but analyze in both cases also the secondary performance measure.

4 Methodology and Data

4.1 Terminal-and-Service Selection Problem (TSSP)

We search here for the path of a single freight shipment with origin o, destination d and departure time $\bar{t}$ through a network with multiple time-sensitive arcs. We model this as a TSSP. For this, we define with $\mathcal{N}$ the set of nodes in the network, which consist of transshipment terminals $\mathcal{T}$, the order's origin o and the order's destination d. Each node $i \in \mathcal{N}$ is associated with emissions e_i^{tship} that are caused by transshipping the order at this facility and time t_i^{tship} that is required for the transshipment of the considered order at node i. With $\mathcal{S}$ we denote the set of services, which include road services and rail services. Here, for a service $s \in \mathcal{S}$, $o_s \in \mathcal{N}$ is the origin node, $d_s \in \mathcal{N}$ is the destination node, e_s^{trans} are the emissions emitted from transporting the order from o_s to d_s using service s, t_s^{dep} is the scheduled departure time at node o_s, t_s^{dur} is the transit time and $suc_s \in \mathcal{S}$ the succeeding service of service s that is conducted by the particular vehicle that performs service s. We use suc_s to model trains with multiple stops, as each link between two stops is modeled as a service on its own. Transshipment processes between a service s and a subsequent service s' are only required if s' is not the succeeding service $suc_s \in \mathcal{S}$ of the same train, i.e. if $s' \neq suc_s$. Note that we set $t_s^{dep} = \bar{t}$ for all services $s \in \mathcal{S}$ where $o_s = o$, which means that the departure times of truck services that depart at origin o is set to the orders departure time. Binary decision variable x_s is equal to 1 if service $s \in \mathcal{S}$ is used in the routing and variable y_s is equal to 1 if a transshipment is required at node $d_s \in \mathcal{N}$. Decision variable t_i states the "ready-for-departure" time at node $i \in \mathcal{N}$, which is the arrival time of a service s at node $d_s = i$ plus the transshipment time t_i^{tship} if transshipment is required at node i. Table 2 provides an overview of sets, parameters and decision variables. The mathematical formulation of the mixed integer linear program is as follows.

$$\min \; t_d - \bar{t} \tag{1}$$

$$\min \; \sum_{s \in \mathcal{S}} (e_s^{trans} \cdot x_s) + \sum_{s \in \mathcal{S}:\; d_s \neq d} (e_{d_s}^{tship} \cdot y_s) \tag{2}$$

subject to

$$\sum_{s \in \mathcal{S}:\; o_s = o} x_s = \sum_{s \in \mathcal{S}:\; d_s = d} x_s = 1 \tag{3}$$

$$\sum_{s \in \mathcal{S}: o_s = d} x_s = \sum_{s \in \mathcal{S}: d_s = o} x_s = 0 \tag{4}$$

$$\sum_{s \in \mathcal{S}: o_s = i} x_s = \sum_{s \in \mathcal{S}: d_s = i} x_s \leq 1, \; \forall \; i \in \mathcal{T} \tag{5}$$

$$(x_s = 0) \implies (y_s = 0), \; \forall s \in \mathcal{S} \tag{6}$$

$$(x_s = 1) \implies (y_s = 1 - x_{suc_s}), \; \forall s \in \mathcal{S} \tag{7}$$

$$t_i = \sum_{s \in \mathcal{S}: d_s = i} x_s \cdot (t_s^{dep} + t_s^{dur}) + y_s \cdot t_{d_s}^{tship}, \; \forall \; i \in \mathcal{T} \tag{8}$$

$$(x_s = 1) \implies (t_{o_s} \leq t_s^{dep}), \; \forall \; s \in \mathcal{S}, \; d_s \neq d \tag{9}$$

$$(x_s = 1) \implies (t_{o_s} + t_s^{dur} = t_d), \; \forall \; s \in \mathcal{S}, \; d_s = d \tag{10}$$

$$x_s, \; y_s \in \{0, 1\}, \; \forall \; s \in \mathcal{S} \tag{11}$$

$$t_i \geq \bar{t}, \; \forall \; i \in \mathcal{N} \tag{12}$$

Table 2. Sets, parameters and decision variables in the TSSP.

Name	Description
$\mathcal{N}$	set of nodes
$\mathcal{T}$	set of transshipment terminals
$\mathcal{S}$	set of services
o	origin of the order
d	destination of the order
$\bar{t}$	departure time of the order
o_s	origin of service $s \in \mathcal{S}$
d_s	destination of service $s \in \mathcal{S}$
e_s^{trans}	emissions caused by using service $s \in \mathcal{S}$ for shipping the order
t_s^{dep}	scheduled departure time of service $s \in \mathcal{S}$ at node o_s
t_s^{dur}	transit time of service $s \in \mathcal{S}$
suc_s	succeeding service ($suc_s \in \mathcal{S}$) of rail service $s \in \mathcal{S}$
e_i^{tship}	emissions caused by a transshipment process at node $i \in \mathcal{N}$
t_i^{tship}	required time for a transshipment process at node $i \in \mathcal{N}$
x_s	$= 1$ if service $s \in \mathcal{S}$ is used
y_s	$= 1$ if transshipment at the end of service $s \in \mathcal{S}$ (at node d_s) is required
t_i	ready-for-departure time at node $i \in \mathcal{N}$

Objective (1) finds the path with the shortest transit time and Objective (2) finds the path that emits the lowest amount of greenhouse gases. Constraints (3) state that the order has to leave the origin and that the order has to arrive at the destination. Constraints (4) ensure that the routing of the order does not use services that end at the origin o or that start at the destination d, which guarantees a cycle-free routing. The balance of flow for the rail stations is described by Constraints (5). The binary transshipment variable is set by Constraints (6) and (7). Constraints (8) to (10) state time-related restrictions. Finally, the domains of the decision variables are defined in Constraints (11) and (12). We implement this TSSP formulation with Version 12.7.1.0 of the IBM ILOG CPLEX Optimization Studio software and solve the problem with a time limit of 300 seconds per order.

4.2 European Road and Rail Network Data

We solve the TSSP repeatably for a large amount of randomly drawn orders across continental Europe where each order has a load of 18 tons (i.e. about one truckload). Regions that can only be reached via ferries, tunnels or large road-detours (e.g. Sicily, Norway, Ireland) and regions that are located east of longitude 19° East are excluded (e.g. Greece, Bulgaria, Romania). Overall, we consider 6,213 randomly created locations as potential origins and destinations of the orders, see Fig. 2 (left).

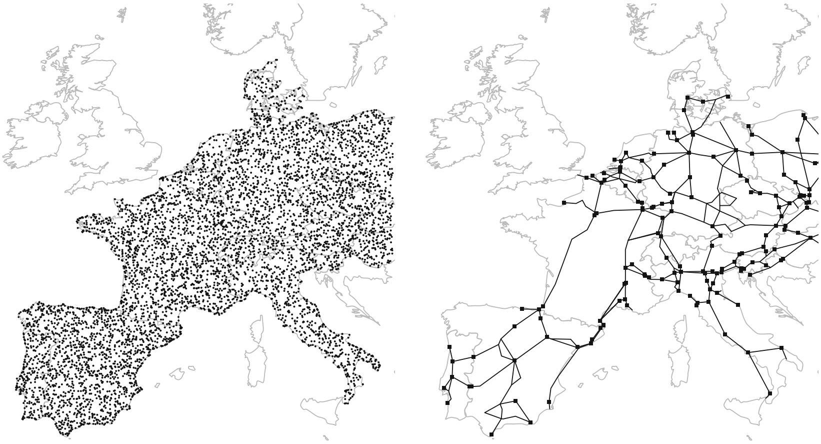


Fig. 2. Randomly created locations (left) and TEN-T railway corridors (right) for continental Europe.

The infrastructural rail network data is based on maps from the Trans-European Transport Network which are provided by the European Commission, see TEN-T Compliance Maps (2014) and TEN-T Interactive Maps (2019). If these maps indicate the traction type of a rail section as electric, we consider the country specific electric power mix as described in Moro and Lonza (2018) for computing the emissions of trains. Otherwise, we assume diesel traction for a train. For the time oriented routing, we use timetables provided at the TEN-T Customer Information Platform, which are available for eight core corridors, see PaP Catalogue (2019). The rail services considered in this study are thus limited to these corridors, see Fig. 2 (right). We assume that every stop mentioned in the timetables qualifies for rail/road transshipment. We combine the information from the timetables with distances from Raildar (2019) to calculate the average speed of trains. For some rail services, this can lead to very high or very low values, for example if the timetables do not mention waiting times at intermediate stops. Therefore, we require the speed to be at least 40 km/h and at most 120 km/h for such services. We model rail services for one full week, which is sufficient to reach every potential destination from every potential origin. For the road network, we use raw data from Open Street Map (2019), as is provided by Geofabrik (2019), and calculate travel time and distance with the Open Source Routing Machine, see Luxen and Vetter (2011). The calculated time is the pure driving time and, thus, does not conform to regulations about drivers' working hours, which may vary between countries. As a general rule the "daily driving time shall not exceed nine hours", see Article 6 in EC Regulation 561 (2006). Therefore, for reasons of simplicity, we do not include break times, if a truck service takes at most 9 h. However, for truck services where the driving time exceeds 9 h, we add another 15 h break time for each full 9 h of driving time.

Additional information on the processing of the road and rail network data can be found in Heinold and Meisel (2018).

For estimating emissions from road and rail transportation, several models were developed, see Hickman et al. (1999), Lindgreen and Sorenson (2005), Scora and Barth (2006) or Wang (1999). In these studies, emissions usually refer to carbon dioxide equivalents (CO_2e) to convert all kinds of greenhouse gases that occur in vehicle operations into CO_2-equivalents. We use the mesoscopic emission estimation model from Kirschstein and Meisel (2015) to estimate emissions from road and rail transportation in the TSSP. The mesoscopic model allows an individual setting of model parameters while keeping the computational effort at a moderate level. Table 3 provides an overview of relevant parameters and their data sources. Note that the payload set for the trains corresponds to a utilization rate of around 70%. Although this rate could be varied, it was shown in Heinold and Meisel (2018) that its impact on the emission rate of an order is relatively low. For each transshipment operation we assume a processing time of 30 min and emissions of 20 kgCO_2e for all terminals $\mathcal{T}$, see Geerlings et al. (2011) and Clausen et al. (2013). For additional information on the mesoscopic model, we refer to Kirschstein and Meisel (2015), Behnke and Kirschstein (2017) and Heinold and Meisel (2018).

5 Simulation

5.1 Generation of Orders

We model intermodal transportation as a three-leg transport chain, consisting of rail transportation between two selected rail/road transshipment terminals (which might consist of multiple consecutive rail services) as well as a road service from the orders origin o to a starting rail terminal and a road service from the last rail terminal to destination d. Thus, we disregard truck transportation between rail services (e.g. road-rail-road-rail-road), although the model presented

Table 3. Parameter values for the emission estimation.

Parameter	Truck	Train[a]
Acc.[b]	0.12	0.05
Distance	Open Source Routing Machine	Raildar (2019)
Gradient[c]	Open Elevation (2019)	Open Elevation (2019)
Payload	18 tons	646 tons
Rail cars	n/a	24 rail cars
Speed	Open Source Routing Machine	PaP Catalogue (2019)[d]

[a]these values are identical for both electric and diesel trains
[b]number of acceleration processes per kilometer
[c]adjusted gradient with $\alpha = 0.5$, see Heinold and Meisel (2018)
[d]combined with distance from Raildar (2019)

in Sect. 4.1 could cover such transport chains too. We consider the twenty closest terminals to the origin and destination as candidates for rail/road transshipment and leave it to the optimization to find the most appropriate transshipment terminals for each order.

Each order is associated with a randomly drawn origin and destination. To simulate orders that are relevant for intermodal transportation, we require the air-line distance between origin and destination to be at least 300 km. The departure time $\bar{t}$ of each order is a randomly generated timestamp in minutes between $[1, 2, ..., 10\ 079]$, where 1 stands for Monday at 0:01 and 10 079 for Sunday at 23:59. If an order uses a rail service $s \in \mathcal{S}$, we apply a payload based allocation scheme and allocate $18t/646t = 2.8\%$ of the total emissions $e^{trans}(s)$ that are caused by the whole train to the considered order.

We use the convergence index from Jung et al. (2004) to find a sufficient number of simulated orders for a reliable estimation of emission rates and transit times, see Fig. 3. We see that 2000 orders lead to negligibly low variations in the average estimated emissions per ton-km and in the average estimated transit time per km. Therefore, all results reported in the following stem from a simulation of 2000 transport orders.

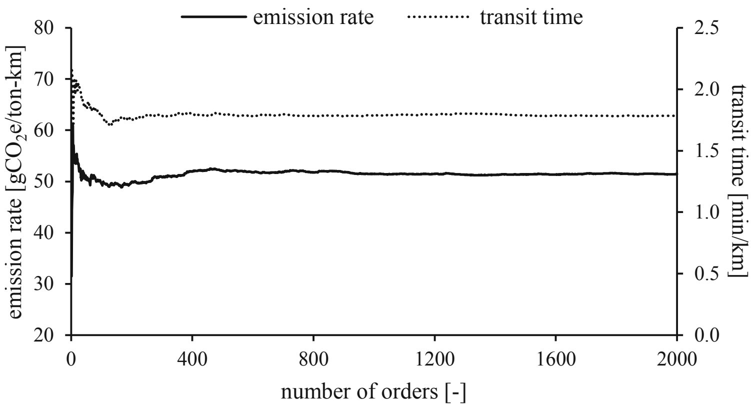


Fig. 3. Convergence of the average emission rate and transit time.

5.2 Results

For all simulated orders, we compare the emission oriented routing and the time oriented routing. Table 4 presents relevant measures from both routings. As expected, the average emissions (in $kgCO_2e$) are lower and the average transit times (in h) are higher for orders that are routed with an emission oriented objective. More precisely, in the time oriented routing, orders have an average transit time of 35 h and arrive at the destination on average 65 h earlier compared to the emission oriented routing. This decrease in transit time corresponds to an average increase of emitted greenhouse gases of 563 $kgCO_2e$ per order. In other words, orders that are routed with an emission oriented objective require on average 890 $kgCO_2e$ which is 40% less than in the time oriented route, while the average transit time almost triples to 101 h on average.

Table 4. Results obtained from the simulation experiments.

	Emission oriented	Time oriented	△
Avg. emissions [$kgCO_2e$]	890.1	1,453.0	+562.9
Avg. emission rate [gCO_2e/ton-km]	51.5	81.7	+30.2
Avg. transit time [h]	100.8	35.5	−-65.3
Avg. transit time [min per km]	5.9	1.8	−4.1
Avg. truck dist [km]	354	859	+505
Intermodal route [%]	97.8	73.8	−24.0
Same routing [%]	2.0	2.0	0.0

Remember that the distance between origin o and destination d is at least 300 km for all simulated orders. However, despite of this requirement, not all orders use rail transportation in the optimal solution. Around 2% of the orders in the emission oriented routing and around 26% of the orders in the time oriented routing use road services exclusively. These road-road routings can only happen for orders where potential transshipment terminals at the origin area overlap with potential transshipment terminals at the destination area. If this is the case, the use of rail services could involve detours that result in higher emissions and/or higher transit times compared to not using rail services at all. However, the majority of the considered orders uses at least one rail service in the optimal solution. For these orders, time oriented routings often select transshipment terminals where the orders' waiting time for the next rail service is low and where well-connected rail services can be used for the rail routing. This practice can lead to longer distances for truck transportation in the time oriented routing. The average truck distance from pre- and post-carriage is 859 km in the time oriented routing whereas it is only 354 km in the emission oriented routing. Eventually, for 2.0% of all orders, the time oriented routing uses the same services as the emission oriented routing. In other words, the chosen objective impacts the selected route for the vast majority of 98% of all orders.

Figure 4 shows the emission oriented and the time oriented routing for an exemplary order from France to Germany. The weekdays and times in the graph refer to the departure at the nodes. In the emission oriented routing, transshipment terminals are selected that are close to the origin and destination and the rail transportation involves a detour via Belgium and the Netherlands. In contrast, the time oriented routing selects more distant terminals, which leads to rail transportation with much lower detouring. In addition, the used rail services

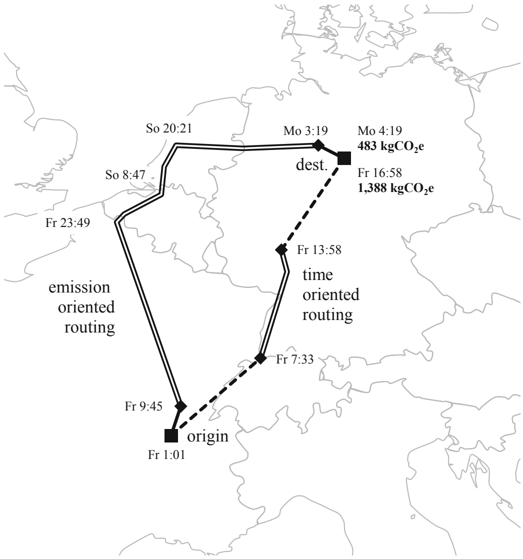


Fig. 4. Example of an emission oriented and a time oriented routing from France to Germany.

in the time oriented routing are well synchronized and incur only little waiting time at rail stations while the used rail services in the emission oriented routing can imply long waiting times between two rail services.

The total emissions and the total transit time increase with the distance between origin o and destination d. To partly control for this, the environmental impact can be measured as gCO_2e/ton-km and the transit time as min/km. In Fig. 5 we put the (o, d)-airline distance of the orders in relation to the emission rate in the emission oriented routing (left graph) and in relation to the transit time per km in the time oriented routing (right graph).

For the emission rate (gCO_2e/ton-km) we observe a non-linear decrease with increasing distances. This is because truck transportation between transshipment terminals and origin/destination causes relatively high emissions per ton-km. This amount depreciates with increasing intermodal transport distances of orders, as the use of the eco-friendly rail mode becomes more and more relevant. The transit time (min/km) appears to increase slightly with the orders distance. This can be explained with additional transshipment operations and waiting times at rail stations as increasing distances usually result in using more rail services.

We further specify the range of emission rates and transit times per km in Fig. 6. Here, we depict the cumulative distribution function of emission rates and transit times for both routing options. In general, orders in the emission oriented routing have a lower emission rate (left graph) and a higher transit time (right graph) compared to if they were routed with the objective to minimize

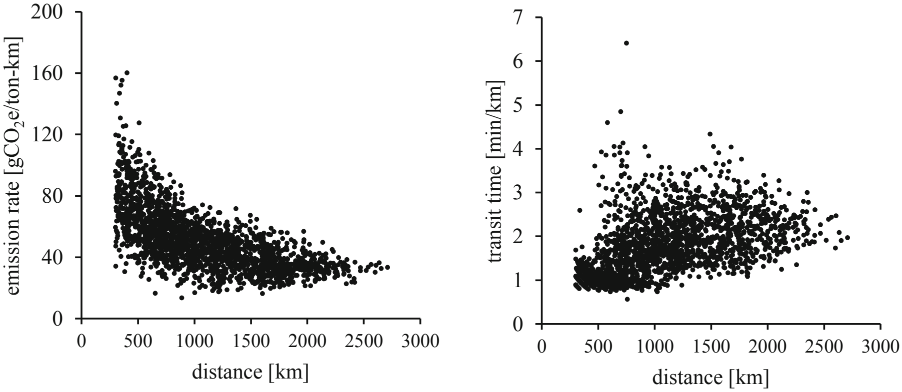


Fig. 5. Relation between emission rates (left), transit times (right) and (o, d)-airline distance.

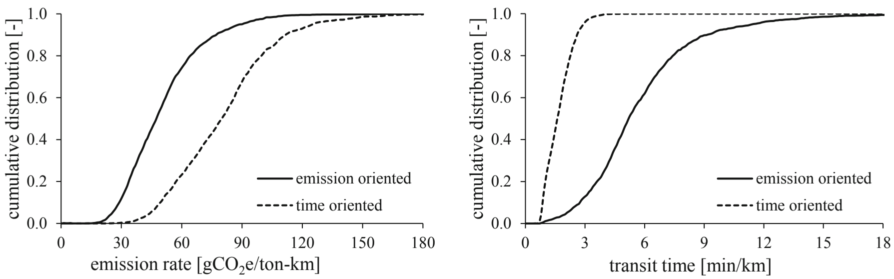


Fig. 6. Cumulative distribution function of emission rates (left) and transit times (right) for both routing options.

time. For example, 50% of the orders (median) have an emission rate of at most 48.4 gCO_2e/ton-km and a transit time of at most 5.3 min/km in the emission oriented routing whereas the emission rate is at most 80.8 gCO_2e/ton-km and the transit time is at most 1.7 min/km in the time oriented routing. The cumulative distribution functions also indicate that the range of optimal values for emission rates and transit times is narrower if the objective includes emissions or time, respectively. In particular, the interquartile range, which indicates the difference between the 75% and the 25% quantile, for emission rates is $61.5 - 36.6 = 24.9$ in the emission oriented routing and $96.0 - 62.6 = 33.4$ in the time oriented routing. The interquartile range for transit times is $7.0 - 4.1 = 2.9$ in the emission oriented routing but only $2.2 - 1.2 = 1.0$ in the time oriented routing.

6 Conclusion

In this paper, we compare the emission oriented routing with the time oriented routing of large freight shipments that have origins and destinations all across Europe. We use the terminal-and-service selection problem (TSSP) to solve the routing problem for each considered order. Our results are based on a simulation of 2000 fixed-sized orders in the European rail/road network. The network consists of realistic timetables for trains and realistic driving times for trucks. We show that substantial differences exist between an emission oriented and a time oriented routing. The average amount of emitted greenhouse gases is 890 $kgCO_2e$ in the emission oriented routing and 1,453 $kgCO_2e$ in the time oriented routing. The average transit time is 101 h in the emission oriented routing and around 36 h in the time oriented routing. The average emission rate is 51 gCO_2e/ton-km in the emission oriented routing and 81 gCO_2e/ton-km in the time oriented routing. The average transit time is 5.9 min/km in the emission oriented routing and 1.8 min/km in the time oriented routing. All results show that the chosen objective has a tremendous impact on the chosen route and its performance measures. We also show that truck services are more relevant in time-oriented routings but that the tightly scheduled rail services of the European TEN-T corridors are often competitive even if transit time is to be minimized. If emissions are to be minimized, intermodal transportation clearly dominates road transportation with substantially lower total emissions for almost all considered orders. Future research can consider the interdependencies of orders that are jointly routed through the network and the impact of stochastic departures and transit times on the reliability of intermodal transport options. Furthermore, it could be interesting to investigate the limitations of the current train network and possibilities to extend it.

Acknowledgments. This research was funded by the German Research Foundation (DFG) under reference ME 3586/1-1. We thank three anonymous reviewers for their valuable comments which helped to improve the manuscript considerably.

References

Bauer, J., Bektaş, T., Crainic, T.G.: Minimizing greenhouse gas emissions in intermodal freight transport: an application to rail service design. J. Oper. Res. Soc. **61**(3), 530–542 (2010)

Behnke, M., Kirschstein, T.: The impact of path selection on GHG emissions in city logistics. Transp. Res. Part E: Logist. Transp. Rev. **106**, 320–336 (2017)

Bektaş, T., Laporte, G.: The pollution-routing problem. Transp. Res. Part B: Methodol. **45**(8), 1232–1250 (2011)

Clausen, U., Kaffka, J, Déring, L., Ebel, G.: Container based calculation of greenhouse gas emissions – a method to determine emissions of container handlings in container terminals. In: General Proceedings of the 13th World Conference on Transport Research (WCTR) (2013)

Crainic, T.G.: Service network design in freight transportation. Eur. J. Oper. Res. **122**(1), 272–288 (2000)

de Miranda Pinto, J.T., Mistage, O., Bilotta, P., Helmers, E.: Road-rail intermodal freight transport as a strategy for climate change mitigation. Environ. Dev. **25**, 100–110 (2018)

Demir, E., Burgholzer, W., Hrušovský, M., Arıkan, E., Jammernegg, W., Van Woensel, T.: A green intermodal service network design problem with travel time uncertainty. Transp. Res. Part B: Methodol. **93**, 789–807 (2016)

Dijkstra, E.W.: A note on two problems in connexion with graphs. Numer. Math. **1**(1), 269–271 (1959)

EC Regulation 561/2006.: Regulation (EC) No 561/2006 of the European Parliament and of the Council of the European Union (2006). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006R0561:EN:HTML. Accessed 22 Jan 2019

Ehmke, J.F., Campbell, A.M., Thomas, B.W.: Data-driven approaches for emissions-minimized paths in urban areas. Comput. Oper. Res. **67**, 34–47 (2016)

Figliozzi, M.: Vehicle routing problem for emissions minimization. Transp. Res. Rec. J. Transpo. Res. Board **2197**, 1–7 (2010)

Floyd, R.W.: Algorithm 97: shortest path. Commun. ACM **5**(6), 345 (1962)

Geerlings, H., van Duin, R.: A new method for assessing CO2-emissions from container terminals: a promising approach applied in Rotterdam. J. Clean. Prod. **19**(6–7), 657–666 (2011)

Geofabrik.: OpenStretMap Data Extracts (2019). https://download.geofabrik.de/. Accessed 22 Jan 2019

Heinold, A., Meisel, F.: Emission rates of intermodal rail/road and road-only transportation in Europe: a comprehensive simulation study. Transp. Res. Part D: Transp. Environ. **65**, 421–437 (2018)

Hickman, J., Hassel, D., Joumard, R., Samaras, Z., Sorenson, S.: Methodology for calculating transport emissions and energy consumption (1999). https://trimis.ec.europa.eu/sites/default/files/project/documents/meet.pdf. Accessed 22 Jan 2019

Jung, J.Y., Blau, G., Pekny, J.F., Reklaitis, G.V., Eversdyk, D.: A simulation based optimization approach to supply chain management under demand uncertainty. Comput. Chem. Eng. **28**(10), 2087–2106 (2004)

Kim, N.S., van Wee, B.: Toward a better methodology for assessing CO_2 emissions for intermodal and truck-only freight systems: a European case study. Int. J. Sustain. Transp. **8**, 177–201 (2014)

Kirschstein, T., Meisel, F.: GHG-emission models for assessing the eco-friendliness of road and rail freight transports. Transp. Res. Part B: Methodol. **73**, 13–33 (2015)

Lam, J.S.L., Gu, Y.: A market-oriented approach for intermodal network optimisation meeting cost, time and environmental requirements. Int. J. Prod. Econ. **171**, 266–274 (2016)

Lin, C., Choy, K.L., Ho, G.T., Chung, S.H., Lam, H.Y.: Survey of green vehicle routing problem: past and future trends. Expert Syst. Appl. **41**(4), 1118–1138 (2014)

Lindgreen, E.B.G., Sorenson, S.C.: Simulation of energy consumption and emissions from rail traffic. Technical report. Technical University of Denmark (2005). http://orbit.dtu.dk/files/3413826/lindgreen_sorenson.pdf. Accessed 22 Jan 2019

Luxen, D., Vetter, C.: Real-time routing with OpenStreetMap data. In: Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 513–516 (2011)

Moro, A., Lonza, L.: Electricity carbon intensity in european member states: impacts on GHG emissions of electric vehicles. Transp. Res. Part D: Transp. Environ. **64**, 5–14 (2018)

Open Elevation: Open Elevation: a free and open-source elevation API (2019). https://open-elevation.com. Accessed 17 Jan 2019

Open Street Map: OpenStreetMap is the free wiki world map (2019). https://www.openstreetmap.org. Accessed 17 Jan 2019

PaP Catalogue: Annual catalogue of the pre-arranged paths (PaP) 2019 (2019). https://cip.rne.eu/apex/f?p=212:170:2793393320158::::P170_BOOKS_LIST:504556. Accessed 22 Jan 2019

Raildar: raildar.fr (2019). http://raildar.fr/osrm/osrm.html. Accessed 22 Jan 2019

Scora, G., Barth, M.: Comprehensive modal emissions model (CMEM), version 3.01 (Users guide). Centre for Environmental Research and Technology. University of California, Riverside (2006). https://www.cert.ucr.edu/cmem/docs/CMEM_User_Guide_v3.01d.pdf. Accessed 22 Jan 2019

TEN-T Compliance Maps: TEN-T Compliance Maps. TENtec Reporting. European Commission (2014). https://ec.europa.eu/transport/themes/infrastructure/downloads_en. Accessed 21 Jan 2019

TEN-T Interactive Maps: TEN-T Interactive Maps. TENtec Reporting. European Commission (2019). http://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/map/maps.html. Accessed 21 Jan 2019

Wang, M.Q.: Greet 1.5-transportation fuel-cycle model-vol. 1: methodology, development, use, and results. Technical report, Argonne National Lab., IL (US) (1999). https://greet.es.anl.gov/files/20z8ihl0. Accessed 22 Jan 2019

Winebrake, J.J., Corbett, J.J., Falzarano, A., Hawker, J.S., Korfmacher, K., Ketha, S., Zilor, S.: Assessing energy, environmental, and economic tradeoffs in intermodal freight transportation. J. Air Waste Manag. Assoc. **58**(8), 1004–1013 (2008)

DEA Sustainability Evaluation in Automotive Supply Chains

Birte Struve(✉), Timo Christopher Anke, and Matthias Klumpp

Georg-August-University Göttingen,
Wilhelmsplatz 1, 37073 Göttingen, Germany
struve.birte@gmail.com

Abstract. The question of sustainability evaluation in global supply chains is often answered qualitatively with standards and status evaluations. However, dedicated management requires also quantitative approaches to evaluate the existing situation adequately. By applying a quantitative approach, this research work considers the global automotive supply chains with a Data Envelopment Analysis (DEA) sustainability evaluation and matching key performance indicators. In this setting, 13 automotive companies are analyzed for three different years 2015 to 2017. Results show that different OEMs are featuring very distinctive sustainability settings and results – deriving also optimization potential in comparison to the other companies.

Keywords: DEA · Sustainability · Automotive supply chains

1 Introduction

While global growth and modern lifestyle are exerting sustainability pressure on our planet, they get more and more into the focus of political and societal measures [1–5]. Therefore, also firms put constantly more effort into sustainability management topics and concepts, especially with regard to the transport and supply chain context [1, 6–8]. This is a long-term strategic challenge for firms as political and societal requirements change constantly and business cycles are influencing possible and adequate measures [8–10]. Especially in the automotive sector, sustainability is a trend topic: McKinsey reports that the expected sales of the industry are rising up to 6.6 billion US-Dollar in 2030, indicating increasing production volumes [11]. Besides growth, also innovation contributions from the automotive sector are evaluated high [12, 13]. Additionally, ecological and social contributions exist as for example in Germany, this sector employs about 820,000 people in 2017 [14, 15]. Therefore, efficient resource use is important for the automotive industry and for the whole economy [13].

The objective of this research work is to establish a Data Envelopment Analysis (DEA) (for further details see e.g. [16–18]) with applicable key performance indicators regarding the sustainability performance in a triple bottom line (TBL) approach for the automotive industry and test this approach with real life business data from company reports.

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 203–220, 2019.
https://doi.org/10.1007/978-3-030-29821-0_14

Recent research papers have been published focusing on the efficiency of sustainability evaluated with the DEA. However, the majority of sustainability-oriented DEA calculations, with the exception of some authors (see e.g. [8, 19, 20]), focus exclusively on eco-efficiency. Consequently, there is a research deficit in measuring and analyzing the efficiency of the multidimensional construct of sustainability under economic, environmental and social dimensions. Besides sustainability-oriented efficiency evaluations, DEA calculations were conducted often in the context of the automotive industry. In these studies, primarily economic performance indicators are included in the calculation of efficiency values. They therefore represent a company-oriented rather than an environmentally and socially oriented efficiency measurement. To date, no efficiency evaluation of a sustainable automotive supply chain in the sense of a TBL approach has been conducted using the DEA. Thus, this research work is addressing a prominent research gap regarding sustainable supply chain management.

The chapter is structured as follows: Sect. 2 provides a detailed description of required, available and used input and output performance data in automotive supply chains. Section 3 is presenting the DEA efficiency analysis results in three consecutive runs with their specific data and settings. Section 4 provides a discussion of the results regarding implications, limitations and comparisons to other research results. Section 5 finally outlines a conclusion and outlook with further research questions for future analysis in this important segment of sustainable automotive supply chain management.

2 Key Performance Indicators

2.1 Requirements for Data Quality

The sustainability-oriented input and output factors, as the target contributions to the efficiency value, must meet different data quality criteria. The key performance indicators need to be critical for the efficiency of the company and the key performance indicators need to be within their control [18, 21]. Furthermore, a certain degree of homogeneity with regard to the environment of the Decision Making Units (DMUs) and its input and output factors used in the production process is required [22, 23]. Besides that, it must be ensured that the values of an input and output are scaled in the same units, e.g. monetary units [24]. Moreover, especially when using environmental factors, no relative performance indicators should be used as they do not allow conclusions about the actual environmental impact when changing the relative value [25]. Another criterion is an appropriate level of input and output indicators used in relation to the number of DMUs. Bowlin [22] recommends that three times the number of variables are required for each input or output used in order to ensure data quality.[1] Lastly, the non-negativity of the individual values need to be guaranteed [22, 24]. In this research work, we succeed in meeting all data quality requirements.

[1] This corresponds to the following calculation: $39 \geq (4 + 5) \times 3$.

2.2 Selection of Decision Making Units

The potential DMUs of this research include any global automotive manufacturers worldwide. We examine the corporate groups and not the individual brands in order to consider possible efficiency improvements resulting from the consolidation of different brands (e.g. synergies and economies of scale). Hence, in the case of groups comprising several different automotive brands, only the consolidated annual, sustainability and financial reports of the groups are considered. The specific selection of DMUs refers to the 20 largest car manufacturers in the world as measured by revenue. However, only automotive companies are considered that generate at least 50% of their revenues through the production of passenger cars. In order to enable a comparison among the corporate groups from 2015 to 2017, the following companies where no longer taken into account, since they did not publish any data for at least one variable for two years or more: SAIC Motor, Chongqing Changan Automobile Company, Great Wall Motors Company Limited, Mazda Motor Corporation, BAIC Group, Geely and Mitsubishi Motors. A representative sample of the automotive companies can be assumed, as the companies considered in this research work represent 75% of the total sales volume with respect to the number of sold cars of the global automotive industry [26], as illustrated in Fig. 1.

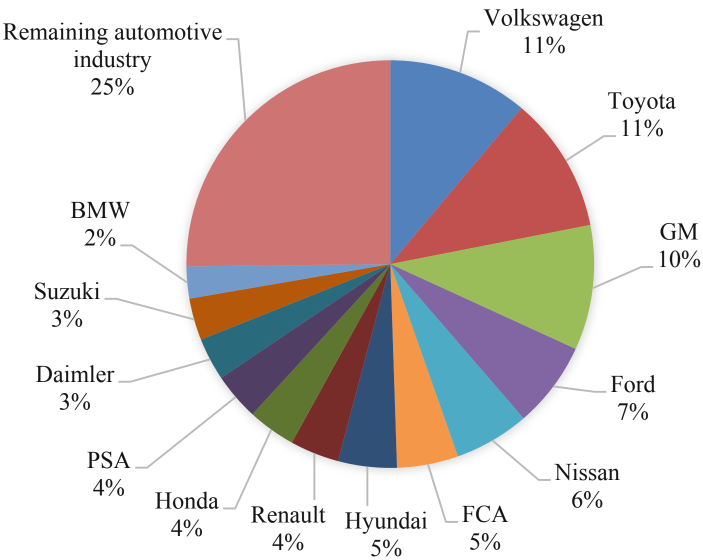


Fig. 1. Share of global sales volume in the automotive industry

2.3 Specification of the Model

In our research work, the software BANXIA Frontier Analyst Version 4 was used for DEA calculations. Table 1 (Annex) illustrates considered sustainability-oriented variables of this research work and shows the classification of inputs and outputs with corresponding sustainability dimensions. The explanations are based on a computational perspective. When defining the key performance indicators, it was ensured that sustainability was comprehensively considered in accordance with the TBL approach. The TBL approach comprises the simultaneous consideration of environmental, social and economic objectives.

When evaluating efficiency in the environmental context, it is necessary to reconsider the classification of indicators in inputs and outputs. In an environmental analysis a distinction must be made between desirable and undesirable outputs. For example, emissions and waste need to be considered as undesirable co-products in the automotive supply chain. In contrast to the desired target contributions, they must be kept as low as possible. One method frequently used in the literature to deal with this problem is that undesirable outputs are used as inputs in DEA calculations and thus being minimized [5, 20, 27–29]. Furthermore, employees represent an input in the production process as a human factor, but the number of total and female employees needs to be maximized according to social aspects (see Table 1 in the Annex). Analogously to the previous argumentation, both key performance indicators of the social dimension need to be classified as outputs. The same principle applies to the key performance indicator equity and research and development (R&D) expenditure. The higher the R&D expenditure spent on improving the efficiency of production and supply chain processes from a social, economic and environmental point of view, the more comprehensively sustainability is addressed. Therefore, R&D expenditure are classified as output and thus being maximized.

In the following, it is necessary to determine which model assumptions correspond the best to the conditions of sustainable automobile supply chain, i.e. whether the input and output variables are either assumed with variable or with constant returns to scale. The CCR model named after Charnes, Cooper and Rhodes [16] is based on constant returns to scale with proportional growth of outputs in relation to inputs. Consequently, the production frontier function is linear. In contrast, the BCC model named after Banker, Charnes and Cooper [17] is based on variable returns to scale. As a result, assuming a BCC model, as many or more DMUs as under a CCR model are considered efficient. As Lozano et al. [30] conclude, the criteria for assuming a CCR or BCC model is not "clear-cut" definable. Banker [31] argues, if all DMUs operate in a competitive market, it can be assumed that each of them operates with the most productive scale size. Consequently, if increasing economies of scale (i.e. variable returns to scale) prevail in a process, every rationally acting company would increase its inputs in order to increase efficiency. Thus, if economies of scale were decreasing, the inputs would be reduced. Hence, if a company operates with its most productive scale size, there are necessarily constant returns to scale. However, this approach by Banker [31] is very generic, but the following DEA is calculated on the assumption of constant returns to scale, as the automotive industry is undoubtedly a competitive market.

Furthermore, it remains to be examined whether an input-oriented or an output-oriented model is appropriate as a model assumption. For this purpose, the production process in the automotive industry is understood as a throughput process. Output orientation defines a model that attempts to maximize output at a given input. Accordingly, an increase in output leads to an increased efficiency score. It is assumed that the primary objective of an automotive company is efficiency in terms of maximizing returns. Thus, it is essential to achieve a maximum of outputs with a given input. Consequently, an output-oriented DEA model is assumed for further calculations.

2.4 Adjustments and Conversions of Values

We collected the data from the annual, sustainability and financial reports communicated by the companies on an annual basis. Since there are no globally obligatory requirements for the publication of these documents [4], differences in the quality and quantity of the reporting of the DMUs exist. Consequently, specific data values of certain DMUs had to be adjusted. All monetary values had to be converted into euros using the exchange rate published by Börse Online [32] in order to ensure uniformity within the economic indicators. Further adjustments were applied to convert the given value of a key performance indicator per vehicle into the total amount of the respective indicator throughout the company. Due to insufficient data on various indicators of the social and environmental dimension, missing values were adjusted. For example, in certain cases the proportion of women in the company was only given for specific countries, hence, it had to be extrapolated to the company's global workforce. If a company was lacking only one value from one of the years 2015, 2016 and 2017, the missing value was adjusted using the average of the other two years. Table 2 in the Annex provides an overview of the data used, presenting minimum, maximum and average values as well as the standard deviation for each input and output type incorporated.

The correlation test revealed a highly positive correlation ($r = 0.96$) between the key performance indicators energy and greenhouse gas (GHG). Such a strong positive correlation can indicate that both variables describe the same phenomenon. Basically, indicators are critical if the efficiency score of a DMU react sensitively to their exclusion [21]. Hence, one indicator was excluded from the model for testing purposes. Although both indicators correlate highly positive with each other, the calculated efficiency scores differ considerably if one of the two factors is excluded from the analysis. For this reason, despite their high correlation, both energy and GHG are kept in the DEA model in order to ensure a larger explanatory value.

3 Sustainability Data Envelopment Analysis

3.1 Analysis 1

The calculated efficiency values of the DMUs as well as the descriptive statistics of the first analysis were compiled in Table 3 (Annex). An efficiency measurement using the DEA does not only provide a relative efficiency value for each unit, but also indicates by how much and where an inefficient unit needs to improve in order to be efficient. Based on this, specific targets can be set for DMUs to improve performance. The calculation results show that the indicators equity, R&D expenditures and number of female employees have the highest average potential for improvements of the DMUs identified by the DEA. The environmental dimension, with its four previously defined key performance indicators, has the lowest potential for improvements. The detailed results of the average potential improvements in the automotive industry are illustrated in Fig. 2.

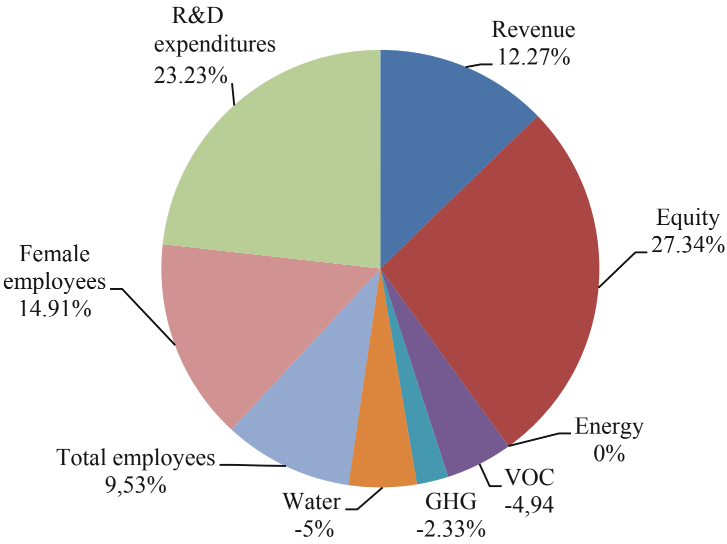


Fig. 2. Average potential improvements of the DMUs

The amount of equity, an indicator of the economic stability and sustainability of an automotive company, represents the highest potential for improvement of all companies on average. The high potential for improvement can indicate that stability and long-term sustainability are no longer maintained. Despite the high importance of equity, this economic indicator is rarely considered in the DEA. However, non-consideration of this factor distorts the long-term assessment of a company's economic sustainability.

The second largest potential for improvement is the level of R&D expenditures, and thus a variable that influences sustainability of a company in all three dimensions. On average, the R&D expenditure of each DMU needs to be increased by 23.23% in order to be classified as relatively efficient. It indicates that all DMUs invest insufficiently in R&D on average. This is particularly critical in today's world, in which the automotive industry faces major challenges that point the way to the future, e.g. electric and autonomous driving as well as new types of business models which require a high degree of innovation. Thus, R&D expenditures in the automotive industry fell from 2015 to 2016 [33]. If the trend of decreasing R&D expenditures continues, the long-term sustainable growth of the automotive industry could be endangered.

The number of female employees as a key performance indicator of the social dimension of sustainability has the third highest potential for improvement of all inputs and outputs considered in the DEA. It is conspicuous that Asian companies in particular have high potential for improvement with regard to the employment of female employees. A country comparison of the proportion of female employees in companies shows that Asian companies have an average proportion of women of less than 15%, whereas American and European car manufacturers have a proportion of women between 15% and 27%. However, an upward trend can be seen in all companies, with the exception of the Japanese company Toyota. Therefore, diversity was valued more highly in the majority of automotive companies.

The consumption of energy and water as well as the emission of GHG and volatile organic compounds (VOC) as environmental sustainability indicators have the weakest potential for improvements. These results indicate that high environmental resource efficiency already prevails in the automotive industry. For all the companies analyzed, great importance is already attached to company-wide environmental programs to reduce consumptions and emissions of these factors. The objectives of these programs were often achieved earlier and exceeded significantly. For example, Ford has already achieved its goal of saving 30% of GHG emissions by 2025, with a saving of 32% already in 2017 [34]. The driving forces behind the implementation of these programs are not only environmental but also economic considerations. For example, a reduced consumption of energy in the automotive production results in significant cost savings.

The aim of this work, as previously defined, is to evaluate the efficiency using a TBL sustainability perspective. Consequently, it is crucial that all three sustainability dimensions are considered in the DEA when calculating efficiency scores. However, for 25 of 39 DMUs only two of the three sustainability dimensions are taken into account. Among them are also efficient DMUs such as Daimler 2015, achieving their efficiency score only on the basis of social and environmental factors. This contradicts the basic idea of the TBL approach. The exclusive consideration of social and environmental performance indicators means that economic variables are not considered when measuring efficiency. However, the achievement of a high economic performance is important in order to sustain the company in the long term and to ensure that it can address social and environmental concerns [1].

In addition, it needs to be examined whether the diversity of inputs and outputs is taken into account in the calculation of efficiency scores. Each of the nine defined inputs and outputs represents a critical aspect for evaluating sustainability in the automotive supply chain. However, the DEA results show that on average the efficiency values were calculated on the basis of only 3.5 inputs and outputs. This also contradicts the idea of our research work to map sustainability in automotive supply chain as comprehensively as possible. In the second analysis, weighting ensures that all inputs and outputs as well as the three sustainability dimensions are considered in the DEA.

3.2 Analysis 2

The three dimensions of sustainability are integrated in the second analysis with an equal minimum weight in order to ensure a comprehensive view of sustainability in terms of a TBL approach. A minimum weighting of 15% is chosen for each of the three economic, environmental and social dimensions, distributed proportionally among the indicators of each dimension. The economic key performance indicators equity and revenue are weighted at 7.5%, the environmental indicators energy, water, GHG, VOC at 3.75% and the number of total and female employees as social indicators at 7.5%. R&D expenditure, as previously defined, is an indicator for each of the three sustainability dimensions and therefore included in the DEA with a minimum weighting of 5%. All factors add up to a weighting of 50%. The remaining 50% can be freely distributed according to the algorithm of the DEA. Hence, the second analysis fulfills both the TBL approach and the specific characteristic of a DEA that DMUs can achieve maximum efficiency by pursuing company-specific strategies.

Due to minimum weights for all sustainability performance indicators within the framework of the DEA, the arithmetic mean of the efficiency values for the automotive companies decreased by 13.1% points (62.8% compared to 75.9%). This indicates an improved discrimination level [8]. When setting a minimum weighting for key performance indicators, not only those indicators for which the company has a particularly favorable output/input ratio compared to other companies are considered, but also those that have an output/input ratio that is unfavorable for company's efficiency. This analysis evaluates efficiency that corresponds more closely to reality, as it is, for example, of high importance for environmental sustainability to reduce not only the emission or consumption of one factor, but all factors simultaneously. An isolated consideration of individual indicators in the DEA is therefore only of limited use. By weighting all key performance indicators and considering them accordingly in the efficiency value, a more comprehensive evaluation of the company's activities with regard to a TBL concept is achieved. The more the weighting reduces the average efficiency value, the higher the difference between the output/input relations compared to the rest of the automotive industry considered in the first analysis and the additional relations due to the introduction of the weighting in the second analysis. Volkswagen and BMW are examples of companies that have shown a decline in efficiency value as a result of the introduction of weighting.

With a 40.1% point decline in the efficiency score between 2015 and 2017, Volkswagen shows the strongest negative change compared to its competitors and none of the DMUs of Volkswagen is efficient any more. Furthermore, comparing the considered inputs and outputs used to calculate the efficiency score in the first and second analysis, energy as input and employees as output continue to be considered most strongly. All other sustainability indicators are only taken into account according to minimum weightings. One reason for Volkswagen's loss of efficiency is the equity base, as indicated by the potential for improvement. The DMU Volkswagen 2015 would have had to increase its equity by 1035% in order to be classified as efficient. In terms of the ratio of equity to the four input performance indicators, Volkswagen has one of the worst ratios in the automotive industry. However, unlike in the first analysis, this unfavorable output/input ratio needs to be included in the DEA due to minimum weightings. The equity base is an indicator of financial independence and long-term stability of a company. The low level of Volkswagen's equity over the three years can be particularly critical in a long-term perspective. Hence, by increasing its equity, Volkswagen could ensure financial independence and thus sustainable stability for the company.

Further noticeably findings that do not appear intuitive at first glance refer to the key performance indicators number of employees and energy. Thus, Volkswagen would have to reduce its workforce or increase its energy consumption in order to be relatively efficient. Intuitively, energy consumption should be reduced and the number of employees increased. This is due to the introduction of a minimum weighting, as that the ratio of employees to energy is higher than the ratio of best practice in the automotive industry and therefore orientation towards this benchmark requires a reduction in workforce and increase in energy consumption. However, as Volkswagen has a relatively unfavorable ratio in all other relations and therefore does not represent best practice in the automotive industry, Volkswagen must, in order to adhere to the logic of the DEA, reduce the ratio of employees to energy.

A comparison of the first and second analysis results of Volkswagen reveals an additional finding. In the second analysis, the input VOC was considered to a lower percentage than in the first analysis. Due to the favorable ratio between the number of employees and energy, DEA will try to integrate as much of this relation as possible into the calculation. However, the availability to integrate the ratio has been limited by introducing a minimum weighting. Consequently, the key performance indicator VOC can no longer be taken into account so strongly, which influenced the efficient score negatively.

For BMW, the average efficiency value declined the least compared to other automotive companies after introducing the minimum weights, even though output/input ratios were considered that are not favorable for BMW. Since BMW is a company with the most balanced factor endowment, it can be stated that the more balanced the factor endowment, the lower the decline in the efficiency value of a company. BMW as best practice of the automotive industry can also be demonstrated on the basis of real business processes. In 2006, BMW set itself the goal of reducing resource consumption in production by 45% until 2020. This objective was already exceeded in 2017 by various research projects with an increase in resource efficiency of 53.2%. Furthermore, a 59% reduction in VOC emissions was achieved between 2006 and 2017. In addition, cost savings of 161 million euros were achieved by reducing the use of resources. BMW is striving for a leading role in the use of renewable energies. In 2017, 81% of the company's electricity already was generated from renewable energy sources. In fact, BMW is one of the most sustainable companies in the automotive industry which can also be seen that it is the only company that has been listed continuously in the Dow Jones Sustainability Index since it was established in 1999 [35, 36].

3.3 Analysis 3

In the two previous analyses, a constant returns to scale CCR model was applied. The choice of an appropriate model, i.e. CCR or BCC model, depends on the process being analyzed. This analysis examines whether a BCC model better reflects a sustainable production process in the automotive industry than the previously assumed constant returns to scale in the form of a single linear efficiency frontier. The same weightings are used as before.

The calculation results of the DEA applying a BCC model deviate from the previous analyses (see Table 3). CCR model calculations show lower efficiency levels than the BCC model calculations. This is due to more restrictive convexity assumptions in the constant returns to scale model. As differences between the calculations occur, the variable returns to scale model indicates a more realistic evaluation of a sustainable automotive supply chain [37, 38]. Accordingly, this seems more appropriate for further analysis as automotive companies are able to exploit economies of scale not proportionally but only to a certain extent depending on their actual size. This can be evaluated by the measure of scale efficiency, i.e. the quotient of the respective efficiency scores of a CCR and BCC model calculations [39].

The values can be between 0 and 1 and are listed in the right column of Table 3 (Annex). A value equal to 1 represents scale efficiency of an individual DMU. In order to analyze whether the reasons for inefficiency is due to a potential scale inefficiency, a distinction needs to be made between pure technical inefficiency, which is associated with operational business, and economies of scale in terms of a suboptimal company size [39]. These two sources of inefficiency can be operationalized using the following formula [39]:

$$[\text{Technical efficiency}] = [\text{Pure technical efficiency}] \times [\text{Economies of scale}] \qquad (1)$$

The technical efficiency represents the calculated efficiency score using a CCR model, as it does not distinguish between economies of scale and pure technical efficiency. On the other hand, BCC model calculations express pure technical efficiency score and, in combination with the results of the CCR model, allow conclusions about scale inefficiencies to interpret the calculated DEA results [39].

For example, the DMU BMW 2017 has a scale efficiency of 1. This means that BMW 2017 is operating in the most productive scale size [31]. This scale efficiency results in a technical efficiency of 1. Thus, it can be concluded that BMW 2017 has both an optimal operating size and an optimal resource allocation and can therefore be used as an example of best practice. The low scores of scale efficiency of Suzuki, Toyota and Volkswagen over the entire examined period are remarkable. However, all DMUs of these three automotive companies (except Suzuki 2017[2]) achieve pure technical efficiency scores between 98% and 100%. These high values indicate that all three companies perform well in terms of operational business. This means that the inefficiency identified in the analysis under the assumption of constant economies of scale was not caused by a suboptimal allocation of resources. Consequently, inefficiency is caused by the scale size of the companies. It is noticeable that Suzuki is the smallest company and Toyota and Volkswagen are the two largest companies among the defined DMUs.[3]

---

[2] The significant increase in Suzuki's pure technical inefficiency score of almost 50% points in 2017 is due to the decline in the ratio of R&D expenditure and female employees to the four environmental inputs.

[3] The size of the companies was defined on the basis of the sales revenues.

This supports the assumption that existing scale inefficiencies result from suboptimal company sizes of the corresponding DMUs. In the case of Suzuki, this means that the company is too small to achieve higher efficiency scores. In contrast, Volkswagen and Toyota are too large to achieve higher efficiency scores. This becomes particularly obvious when examining the returns to scale of the individual DMUs (see Table 3). For example, Toyota and Volkswagen have decreasing economies of scale when pure technical inefficiencies exist. This indicates that both companies are too large compared to corresponding companies in this industry to generate similarly high output/input ratios as the reference units. On the other hand, Suzuki shows increasing economies of scale in 2017. This means that Suzuki is not yet operating at its most productive scale size, but, as previously explained, is too small to achieve economies of scale and scope.

4 Discussion

This research work has pointed at the research gap of missing efficiency evaluations in automotive supply chains in the context of a TBL sustainability analysis using OR tools such as for example DEA. Accordingly, the question arises why the DEA methodology has not yet been used in such a context. The majority of the studies examine the effectiveness of sustainability management using a regression analysis [40–45].

The main focus of these studies is to measure the relationship between the three dimensions of sustainability and the impact of environmental and social activities on business success. Thus, it needs to be examined further, what the corresponding advantages and disadvantages are in the measurement of sustainability on the basis of the DEA or a regression. A DEA calculation delivers a unique, specific result for each individual DMU, which both identifies the causes of inefficiencies and provides specific benchmarks for reducing existing inefficiencies. The consideration of individual DMUs from real-world practice represents the largest differentiation to regression analyses [22]. However, the DEA does not represent a model for predicting the performance of a company for years that have not been included in the evaluation. This is possible with a regression analysis. Consequently, regression analysis may be useful to estimate or to predict the future performance of an entire group of companies, such as an industry [22]. These differences between the two methods show that both approaches have a legitimate position in research. Finally, the choice of a suitable method depends on the research objective. In this research work, the DEA represents a suitable instrument for measuring the efficiency of automotive supply chains from a sustainability perspective. Thus, the previous research gap cannot be explained by the non-applicability of DEA.

In all three analyses, the DMUs (of the years 2015, 2016 and 2017) of the various companies were included in the DEA. However, this could be seen as critical, as it allows a comparison of different companies from different years, which would not occur in real practice. It is feasible, for example, that an efficient company in 2015 was used as a reference unit for calculating the relative efficiency scores of companies in 2017.

Furthermore, the inputs used in the DEA calculations consist only of environmental key performance indicators. As a result, it is not possible to evaluate efficiency without taking environmental variables into account. Although this corresponds to an integrative TBL sustainability concept, it contradicts the first analysis of this research work, in which the analysis was conducted without weighting. This could have been avoided by also defining social and economic inputs, e.g. cost of sales as an economic indicator and the level of fluctuation as a social indicator. Moreover, no environmental sustainability indicator was defined as an output in the DEA. In this context, it would have been reasonable to use indicators such as recyclable materials or renewable energies.

Moreover, it needs to be questioned whether the chosen indicators reflect sustainable corporate performance comprehensively. For example, in the social sustainability dimension, only quantitative characteristics, such as the number of (female) employees, are described. Qualitative characteristics, such as the promotion of work-life balance, flat hierarchies or more flexible working hours and places are not considered in this DEA.

Furthermore, the classification of a key performance indicator in only one of the three sustainability dimensions needs to be critically questioned. Thus, the increase of the quantitative workforce can have both positive effects on the social dimension and negative effects on the economic dimension in terms of an increase in costs.

5 Conclusion and Outlook

Besides company-specific results, this research has largely provided a special contribution in proving the value of the quantitative efficiency evaluation method of DEA for a comprehensive sustainability analysis of automotive supply chains. This is important as many research directions regarding sustainability concepts are featuring a qualitative setting – and this may require complementary perspectives of data-driven results like in this research setup presented here.

Further research is warranted in the following directions: First, other industries and supply chain may be evaluated with the proposed method and indicators. Second, further method development e.g. with sensitivity analysis might be valuable as cross-chain effects and interdependencies are high especially for sustainability issues. Third, further analyses regarding causal effects and management options derived from that are important in order to provide hints for sustainability improvements in the automotive as well as other industries.

Altogether, comprehensive and quantitative sustainability research is worthwhile further attention from the research as well as business practice side in supply chain management.

Annex

Table 1. Overview of the key performance indicators

	Explanation	Indicator	Literature
Input			
VOC[a]	Organic, varied and ubiquitous chemicals with high vapor pressure that are emitted during painting processes in automobile production	Harmful effects on humans and the environment	
GHG[a]	Gases (like carbon dioxide) causing greenhouse effect emitted in the production process	Atmospheric pollution with strong impact on global climate	[8, 20]
Water[a]	Water consumption from various sources in the production process	Company's sustainable and responsible use of water	[29, 46, 47]
Energy[a]	Energy sources such as electricity, natural gas, district heating in the production process	Company's sustainable and responsible use of energy sources	[20, 29, 46, 48]
Output			
Equity[b]	Funds made available to the company by its owners for an indefinite period of time	Financial independence and long-term stability of a company	[50]
Revenue[b]	Cash equivalents and receivables that the company realized through selling its products	Size of the company	[8, 50]
Total employees[c]	Number of employees in the company	Creating jobs and income for the members of a society to fulfil social responsibility	[8, 20, 50, 51]
Female employees[c]	Number of female employees in all hierarchy levels in the company	Gender equality and diversity within a company	[8]
R&D[a, b, c]	Expenditure for research and development of the company	Investments to increase efficiency through innovation, to protect the environment through new technologies and to improve political and social systems in the company	

[a] Key performance indicator of the environmental sustainability dimension
[b] Key performance indicator of the economic sustainability dimension
[c] Key performance indicator of the social sustainability dimension

Table 2. Overview of data used

	Maximum	Minimum	Arithmetic mean	Standard deviation
VOC	27,948	2,358	10,835	7,326
GHG	9,510,000	1,090,000	3,996,352	2,572,979
Water	53,770,000	4,819,684	25,056,588	13,090,733
Energy	25,888,909	4,840,969	12,741,107	6,672,603
Equity	147,460,705,190	8,105,159,090	43,872,581,552	32,865,959,257
Revenue	232,346,287,305	25,659,966,482	115,499,907,778	58,277,779,116
Total employees	642,292	61,601	214,594	143,043
Female employees	104,693	6,044	35,409	26,405
R&D	13,672,000,000	841,229,513	5,125,273,358	3,382,000,820

Table 3. List of DEA efficiency scores for analyses 1–3

Company	Year	Analysis 1	Analysis 2	Analysis 3	
					Scale efficiency
BMW	**2015**	100	96.3	100 −	0.963
	2016	97.1	94.4	94.5 ↑	0.999
	2017	100	100	100 −	1.000
Daimler	**2015**	100	94.8	97.5 ↓	0.972
	2016	100	97.4	98.9 ↓	0.985
	2017	100	98.8	100 −	0.988
FCA	**2015**	74.1	50.1	62.8 ↓	0.798
	2016	75.6	53.4	65.9 ↓	0.810
	2017	76.9	54.6	68.5 ↓	0.797
Ford	**2015**	74.0	56.6	70.1 ↓	0.807
	2016	79.8	62.3	75.4 ↓	0.826
	2017	82.7	64.4	76.8 ↓	0.839
GM	**2015**	41.8	37.8	56.6 ↓	0.668
	2016	44.4	39.6	62.5 ↓	0.634
	2017	46.2	39.8	57.6 ↓	0.691
Honda	**2015**	66.3	60.3	79.2 ↓	0.761
	2016	65.6	60.9	75.6 ↓	0.806
	2017	64.3	58.9	76.0 ↓	0.775
Hyundai	**2015**	86.4	65.5	67.6 ↑	0.969
	2016	92.3	71.2	74.0 ↑	0.962
	2017	100	79.5	100 −	0.795

(*continued*)

Table 3. (*continued*)

Company	Year	Analysis 1	Analysis 2	Analysis 3	
					Scale efficiency
Nissan	**2015**	64.8	56.9	61.1 ↓	0.931
	2016	58.6	50.6	61.6 ↓	0.821
	2017	60.5	54.3	63.7 ↓	0.852
PSA	**2015**	62.8	48.7	50.2 ↓	0.970
	2016	63.7	50.6	52.4 ↓	0.962
	2017	60.5	54.3	54.5 ↓	0.996
Renault	**2015**	91.9	75.2	75.3 ↓	0.999
	2016	92.5	75.2	75.4 ↓	0.997
	2017	100	92.0	92.6 ↓	0.994
Suzuki	**2015**	55.8	41.9	100 ↑	0.419
	2016	53.5	40.3	100 −	0.403
	2017	52.3	40.0	54.9 ↑	0.729
Toyota	**2015**	59.0	52.9	100 −	0.529
	2016	61.8	54.8	100 −	0.548
	2017	61.6	54.3	98.9 ↓	0.549
Volkswagen	**2015**	100	57.1	98.1 ↓	0.582
	2016	95.3	55.5	98.1 ↓	0.566
	2017	97.1	59.6	100 −	0.596
Arithmetic mean		75.9	62.8	79.4	
Minimum		41.8	37.8	50.2	

References

1. Balderjahn, I.: Nachhaltiges Management und Konsumentenverhalten. UVK Verlagsgesellschaf, Konstanz (2013)
2. Carnau, P.: Nachhaltigkeitsethik: Normativer Gestaltungsansatz für eine global zukunftsfähige Entwicklung in Theorie und Praxis. Rainer Hampp Verlag, München (2011)
3. Kraus, P.: Die Auswirkung von Corporate Governance und Nachhaltigkeit auf den Unternehmenserfolg: Eine Betrachtung im Kontext der wertorientierten Unternehmensführung. EUL Verlag, Lohmar (2011)
4. Binder, U.: Nachhaltige Unternehmensführung: Radikale Strategien für intelligentes, zukunftsfähiges Wirtschaften. Haufe Gruppe, Freiburg (2013)
5. Dyckhoff, H., Allen, K.: Measuring ecological efficiency with data envelopment analysis. Eur. J. Oper. Res. **132**(2), 312–325 (2001)
6. König, M.: Stakeholderorientierte Gestaltung der nachhaltigen Unternehmensführung: Konzeption und Empirie. Verlag Dr. Kovač, Hamburg (2013)
7. Lee, K.H., Farzipoor Saen, R.: Measuring corporate sustainability management: a data envelopment analysis approach. Int. J. Prod. Econ. **140**(1), 219–226 (2012)
8. Klumpp, M.: How to achieve supply chain sustainability efficiently? Taming the triple bottom line split business cycle. Sustainability **10**(2), 1–23 (2018)

9. Altenburger, R.: Gesellschaftliche verantwortung als innovationsquelle. In: Altenburger, R. (ed.) CSR und Innovationsmanagement: Gesellschaftliche Verantwortung als Innovationstreiber und Wettbewerbsvorteil, pp. 1–18. Springer Gabler, Berlin (2013)
10. Hanke, T., Stark, W.: Strategy development: conceptual framework on corporate social responsibility. J. Bus. Ethics **85**(3), 507–516 (2009)
11. OICA World Motor Vehicle Production. http://www.oica.net/category/production-statistics/2017-statistics/. Accessed 16 Jan 2019
12. Wang, C.N., Nguyen, X.T., Wang, Y.H.: Automobile industry strategic alliance partner selection: the application of a hybrid DEA and grey theory model. Sustainability **8**(2), 1–18 (2016)
13. Hutter, C.: Nachhaltigkeitsstrategieentwicklung: Das Spannungsfeld von Unternehmen und Stakeholdern in der automobilen Unternehmenspraxis. Springer Gabler, Wiesbaden (2012)
14. Statista: Beschäftigte in der deutschen Automobilindustrie in den Jahren 2005 bis 2017. https://de.statista.com/statistik/daten/studie/30703/umfrage/beschaeftigtenzahl-in-der-automobilindustrie/. Accessed 18 Jan 2019
15. Statista: Anzahl der Beschäftigten nach Industriezweigen in Deutschland. https://de.statista.com/statistik/daten/studie/158670/umfrage/groesste-arbeitgeber-in-der-industrie-in-deutschland/. Accessed 18 Jan 2019
16. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. Eur. J. Oper. Res. **2**(6), 429–444 (1978)
17. Banker, R.D., Charnes, A., Cooper, W.W.: Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag. Sci. **30**(9), 1078–1092 (1984)
18. Lovell, C.A.K.: Linear programming approaches to the measurement and analysis of productive efficiency. Top **2**(2), 175–248 (1994)
19. Cortes, A.A.: Triple bottom line approach for measuring supply chains sustainability using data envelopment analysis. Eur. J. Sustain. Dev. **6**(3), 119–128 (2017)
20. Sartori, S., Alvarenga, T.H.P., Gibim, C., Campos, L.M.S.: Data Envelopment Analysis in the Sustainability Context - a Study of Brazilian Electricity Sector by Using Global Reporting Initiative Indicators. http://www.advancesincleanerproduction.net/fifth/files/sessoes/5B/6/sartori_et_al_academic.pdf. Accessed 28 Feb 2019
21. Hammerschmidt, M., Wilken, R., Staat, M.: Methoden zur Lösung grundlegender Probleme der Datenqualität in DEA-basierten Effizienzanalysen. Die Betriebswirtschaft **69**(2), 289–309 (2009)
22. Bowlin, W.E.: Measuring performance: an introduction to Data Envelopment Analysis (DEA). J. Cost Anal. **15**(2), 3–27 (1998)
23. Backhaus, K., Wilken, R., Bröker, O., Brüne, P.A., Reichle, F.: Effizienzmessung industrieller Dienstleistungen mittels Data Envelopment Analysis – Projekt ServDEA. In: Möller, K., Schultze, W. (eds.) Produktivität von Dienstleistungen, pp. 53–133. Springer Fachmedien, Wiesbaden (2014)
24. Wilkens, S.: Effizientes Nachhaltigkeitsmanagement. Deutscher Universitäts-Verlag, Wiesbaden (2007)
25. Müller, A.: Nachhaltigkeits-Controlling. Uni-edition, Berlin (2011)
26. OICA World Motor Vehicle Sales. http://www.oica.net/category/sales-statistics/. Accessed 16 Jan 2019
27. Blank, R., Krusenbaum, S., Steven, M.: Effizienzanalyse mittels Data Envelopment Analysis unter Berücksichtigung von Emissionen. Wirtsch. Stud. **44**(11), 613–622 (2015)
28. Kuosmanen, T.: Weak disposability in nonparametric production analysis with undesirable outputs. Am. J. Agric. Econ. **87**(4), 1077–1082 (2005)

29. Zhang, B., Bi, J., Fan, Z., Yuan, Z., Ge, J.: Eco-efficiency analysis of industrial system in China: a data envelopment analysis approach. Ecol. Econ. **68**(1–2), 306–316 (2008)
30. Lozano, S., Iribarren, D., Moreira, M.T., Feijoo, G.: The link between operational efficiency and environmental impacts: a joint application of Life Cycle Assessment and Data Envelopment Analysis. Sci. Total Environ. **407**(5), 1744–1754 (2009)
31. Banker, R.D.: Estimating most productive scale size using data envelopment analysis. Eur. J. Oper. Res. **17**(1), 35–44 (1984)
32. Börse Online: Devisenrechner. https://www.boerse-online.de/devisen/devisenrechner. Accessed 10 Nov 2018
33. PWC: Five trends transforming the Automotive Industry. https://www.pwc.at/de/publikationen/branchen-und-wirtschaftsstudien/eascy-five-trends-transforming-the-automotive-industry_2018.pdf. Accessed 11 Dec 2018
34. Ford: Sustainability Report 2017/18. Ford Motor Company, Dearborn (2018)
35. BMW: BMW Group Geschäftsbericht 2017. BMW Group, München (2018)
36. BMW: Sustainable Value Report 2017. BMW Group, München (2018)
37. Feng, C., Chu, F., Ding, J., Bi, G., Liang, L.: Carbon Emission Abatement (CEA) allocation and compensation schemes based on DEA. Omega **53**(1), 78–89 (2015)
38. Jain, R.K., Natarajan, R.: A DEA study of airlines in India. Asia Pac. Manag. Rev. **20**(4), 285–292 (2015)
39. Cooper, W.W., Seiford, L.M., Tone, K.: Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, 2nd edn. Springer, New York (2007)
40. Glaser, J., Hornung, S.: Nachhaltigkeit und Geschäftserfolg, http://www.oekom-research.de/homepage/german/Studie_TUM_oekom.pdf. Accessed 2 Jan 2019
41. Fiori, G., di Donata, F., Izzo, M.F.: Corporate Social Responsibility and Firms Performance. An Analysis on Italian Listed Companies. https://eprints.luiss.it/374/1/Fiori_2007_03_OPEN.pdf. Accessed 28 Feb 2019
42. Samy, M., Odemilin, G., Bampton, R.: Corporate social responsibility: a strategy for sustainable business success. An analysis of 20 selected British companies. Corp. Gov. **10**(2), 203–217 (2010)
43. Aras, G., Aybars, A., Kutlu, O.: Managing corporate performance: investigating the relationship between corporate social responsibility and financial performance in emerging markets. Int. J. Prod. Perform. Manag. **59**(3), 229–254 (2010)
44. Tsoutsoura, M.: Corporate Social Responsibility and Financial Performance. https://pdfs.semanticscholar.org/ece0/ac42df48ec224765ecf4d65df5fd79b549b5.pdf. Accessed 16 Jan 2019
45. Bassen, A., Hölz, H.M., Schlange, J.: The Influence of Corporate Responsibility on the Cost of Capital. https://www.schlange-co.com/wp-content/uploads/2017/11/SchlangeCo_2006_CostOfCapital.pdf. Accessed 16 Jan 2019
46. Masternak-Janus, A., Rybaczewska-Błazejowska, M.: Comprehensive regional eco-efficiency analysis based on data envelopment analysis: the case of polish regions. J. Ind. Ecol. **21**(1), 180–190 (2017)
47. Iribarren, D., Hospido, A., Moreira, M.T., Feijoo, G.: Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. Sci. Total Environ. **409**(10), 1786–1798 (2011)

48. Liu, W., Tian, J., Chen, L., Lu, W., Gao, Y.: Environmental performance analysis of eco-industrial parks in China: a data envelopment analysis approach. J. Ind. Ecol. **19**(6), 1070–1081 (2015)
49. Suh, Y., Seol, H., Bae, H., Park, Y.: Eco-efficiency based on social performance and its relationship with financial performance: a cross-industry analysis of South Korea. J. Ind. Ecol. **18**(6), 909–919 (2014)
50. Chen, Y.: Productivity of automobile industries using the malmquist index: evidence from the last economic recession. J. Cent. Cathedra **4**(2), 165–181 (2011)
51. Maritz, A., Shieh, C.J.: Performance analysis of automobile industry in taiwan with data envelopment analysis. Int. J. Appl. Math. Stat. **38**(8), 84–95 (2013)

Systematization of Humanitarian NGOs from a Logistical Viewpoint

An Exploratory Study in Germany

Christian Hein(✉)

Chair of Business Management, esp. Logistics,
Technische Universität Dresden, 01069 Dresden, Germany
christian.hein@tu-dresden.de

Abstract. Logistics is an important factor in humanitarian aid operations, whether after the outbreak of a natural disaster or during man-made crises. The speed and effectiveness of the various actors depend, among other things, on the appropriate design of the operations, which is why the scientific interest in this field has increased considerably in recent years. Unfortunately, there is a gap between science and practice [1]. One reason for the lack of applicability in humanitarian logistics research can be seen in the diversity of humanitarian actors. As there is no systematic characterization of humanitarian NGOs regarding aspects relevant from a logistical perspective, this paper aims at providing a suitable overview. A classification scheme is developed that includes the four dimensions of the mandate and sector, as well as the internal network structure and the various forms of on-site cooperation. The results can help to better align scientific research in humanitarian logistics and to foster practical exchange between the pertinent humanitarian organizations.

Keywords: Humanitarian logistics · Disaster management · NGO · Classification · Exploratory study

1 Introduction

No matter the precautions we take, disasters can strike anywhere, uprooting lives and destroying property and infrastructure, leaving lasting consequences long after the disaster itself has subsided. Human conflicts and the resulting suffering have been a commodity of the mass media's reporting for decades. One recent example, the crisis in Syria, has caused more than 5.7 million people to flee the country since the violence broke out in 2011 [2]. In addition to the harm we inflict on each other, natural disasters continue to wreak havoc on communities and people, often with little or no warning at all. During the last two decades, more than 1.3 million people have been killed by natural disasters and a further 4.4 billion were left injured, displaced, or in need of humanitarian assistance [3]. According to UNISDR, the expected losses will increase due to climate change [4]. Consider Indonesia, which was hit by natural disasters repeatedly in 2018, including earthquakes, tsunamis, floods, and landslides [5].

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 221–237, 2019.
https://doi.org/10.1007/978-3-030-29821-0_15

A multitude of humanitarian organizations are on site there in order to alleviate human suffering, support rehabilitation and reconstruction, and foster long-term development. These organizations are part of a worldwide and growing system of humanitarian aid, organizations which share the same basic principles, but can be very diverse in their specific objectives as well as in their way of working, decision-making, and financing [6].

In response to disasters, humanitarian assistance must be provided in a timely manner to support the affected communities. Particularly uncertainties regarding trigger, time, and extent of disasters pose significant challenges to the responders and increase the complexity of humanitarian operations [7]. Moreover, logistics play a crucial role, as it significantly influences the effectiveness of relief and determines the speed, the coverage, and thus the costs of humanitarian organizations [8,9]. The research field of humanitarian logistics has grown considerably in recent years and more and more approaches are being developed to meet the specific challenges the field poses. However, practical applicability of research results remains limited [10]. Kunz et al. identified different reasons for the discrepancy between science and practice, including poor problem definition and low contextualization [1]. The particular characteristics of the humanitarian context are further highlighted by Pedraza-Martinez et al. [11]. As a result, the impact of humanitarian logistics research remains limited.

One reason for the lack of applicability in humanitarian logistics research can be seen in the diversity of humanitarian actors. Solutions are therefore rather specific and difficult to generalize or to transfer. There is no systematic characterization of humanitarian organizations regarding aspects relevant from a logistical perspective. This applies especially to the group of non-governmental organizations (NGOs), which is very diverse and inconsistent and can have worldwide networks that are considerably segmented [12]. Therefore, this paper aims at providing a systematic overview of these humanitarian actors in order to align scientific research and to foster practical exchange between organizations which are more or less similar. The following research questions will be answered:

RQ1: What differences exist between humanitarian NGOs regarding logistical issues?

RQ2: How can these organizations be classified systematically according to aspects relevant to logistics?

To address these questions, 75 German humanitarian organizations have been selected and examined indepth according to logistics related criteria. The remainder of this work is structured as follows: After a short characterization of humanitarian logistics, a brief introduction to the international and German humanitarian system will be provided. In section three, the methodology is outlined including the selection of the organizations analyzed and the development of the related categories. An illustration and discussion of the results will be given in section four. Finally, the conclusion and an outlook are provided.

2 Theoretical Background

2.1 Logistics in Humanitarian Operations

Although logistics has always been involved in humanitarian operations, its status in practice and in the academic literature has only received attention of any significance in the last 20 years [13], which seems to be overdue as it plays a particularly important role. Logistics account for a large part of humanitarian operations and influences its effectiveness significantly [14]. The scope of humanitarian logistics can be very broad, as it is used as "an umbrella term for a mixed array of operations" [8]. Humanitarian logistics definitions resemble those of logistics in a commercial context. Thomas and Kopczak define it as "the process of planning, implementing and controlling the efficient, cost-effective flow and storage of goods and materials, as well as related information, from the point of origin to the point of consumption for the purpose of alleviating the suffering of vulnerable people" [15]. Baumgarten recommends that humanitarian logistics should engage preferably where people cannot support themselves [9]. Hence, two basic views on humanitarian logistics have emerged: logistics in continuous aid work and disaster relief logistics [8]. Humanitarian logistics as continuous aid work thereby refers to countries which are incapable of supporting themselves and can be seen as part of development cooperation, whereas humanitarian logistics in disaster relief is less limited to specific regions [16].

Humanitarian logistics includes a wide range of activities and processes, such as preparedness, procurement, transport, warehousing, and training across all disaster phases [8]. Although these tasks are closely related to those in a commercial context, significant differences characterize the humanitarian environment and pose specific challenges [17]. Humanitarian logistics is shaped by a highly dynamic and volatile environment. As the time, the location, the type, and the size of a disaster is often unpredictable, humanitarian logistics must cope with these uncertainties [18]. In addition, humanitarian logisticians must work under great time pressure; while delays might be acceptable in the commercial context, a timely response in humanitarian operations is essential for people's survival [14]. Beamon and Balcik distinguish between the basic processes of assessment, procurement, and shipping [18], Balcik et al. classify the three main stages of supply acquisition/procurement, pre-positioning/warehousing, and transportation [19]. Hellingrath and Widera developed a process-oriented relief chain to describe the flow of supplies. Initially the supplies have to be procured and then stored at the country of origin. This is followed by the task of transportation to and storage at the point of entry. Finally, the last mile distribution has to be undertaken. In addition, other tasks, such as personnel and information services accompany these processes [20].

2.2 International Humanitarian System

The volume of humanitarian assistance has increased significantly in recent years [21]. In 2017, it was estimated that more than 200 million people worldwide required assistance at some point, most of them concentrating in just a few countries. A total of $27.3 billion were allocated to humanitarian responses, also an increase over recent years. Private donations account for $6.5 billion, the vast majority is provided by governments and EU institutions, whereby a small number of donor governments account for the majority of these public funds [22]. The money is channeled through a variety of different actors, which can be divided into the following three: The Red Cross and Red Crescent movement, UN agencies, and NGOs [23]. The Red Cross inheres a special role within the humanitarian system. With its 190 national societies and the international organization which oversees their activities, the IFRC, as well as the specially mandated ICRC, it describes itself as the world's largest humanitarian network [24]. Key players within the humanitarian system are also the multilateral UN organizations, of which the WFP, UNHCR, and UNICEF are the largest. NGOs can operate on an international, national, and local level, are extremely diverse, and are involved in all phases of the disaster management process. The majority of humanitarian assistance is provided by international organizations, a few of which are much larger and dominate the field. Next to the international players, there exists a vast assortment of national and local NGOs, providing assistance within their national or local borders. They often have a closer point of contact with the affected communities [25].

In Germany, the picture of humanitarian assistance is shaped by various governmental and non-governmental actors. The German response is characterized by a large amount of governmental funds and private donations, particularly in response to large sudden onset disasters. However, slow-onset disasters are unfortunately less considered, often referred to as "forgotten" crises [26]. As for governmental actors in Germany, both the Federal Foreign Office and the Federal Ministry for Economic Cooperation and Development (BMZ) are primarily responsible for humanitarian assistance. The ministries do not perform humanitarian projects on their own, but they work together with partners to implement specific measures. These partners include UN agencies or other international organizations such as the IFRC, but also national partners. The main national partners include the German Corporation for International Cooperation (GIZ) and the Federal Agency for Technical Relief (THW). In addition, the German government supports national NGOs [27]. Several NGOs coordinate themselves under the roof of the Association of German Development and Humanitarian Aid NGOs (VENRO) within a working group [28]. However, a systematic overview of German humanitarian NGOs, particularly regarding aspects relevant from a logistical viewpoint, does not exist yet.

3 Methodology

3.1 Methodical Approach

As little scientific knowledge and no comprehensive systematization of humanitarian actors exists as of yet, an open and explorative approach through qualitative-empirical research methods is most suitable for filling the research gap [29]. A content analysis was carried out based on publicly accessible materials of the organizations, including in most cases their homepages and annual reports. Organizations based in Germany were considered for this study to ensure the desired comparability of the research results. There is no uniform register of all humanitarian organizations in the field of disaster relief and development cooperation. The German Central Institute for Social Issues (DZI) provides comprehensive information about the seriousness of donor organizations and awards a seal of approval [30]. Therefore, their database was consulted to identify the corresponding research objects. At the time of the investigation (January 2019), 75 organizations were classified in the category of disaster relief[1]. This selection does not claim to be exhaustive. However, it includes a variety of different actors of which are almost half also representatives in the German association of NGOs VENRO [28].

The analysis technique used in this paper for data evaluation is based on structuring of content, which seeks to extract and summaries information on specific topics and content areas from the material. Qualitative methods are subject to general criticism in comparison to quantitative research methods. This criticism is particularly directed towards the validity and reliability of the results obtained [31]. In the interest of qualitative research, strict adherence to quality standards should therefore be sought to ensure general recognition of the generated results. Construct validity ensures appropriate measures for answering the research questions [32]. Internal validity ensures the significance of the causal mechanisms under study, which is why the results of data analysis were checked for plausibility with the relevant literature [31]. The procedure of analysis was documented to guarantee repeatability and to ensure reliability [33].

3.2 Development of the Criteria for Analysis

The criteria for analysis were derived from a mixed deductive and inductive approach which is suitable for qualitative data analysis. The main categories were deductively constructed according to the knowledge from literature and inductively extended by corresponding attributes while working through the materials. The documents were processed cyclically, following the recommendation of Mayring [34], whereby a structured adaptation could be carried out. The result of this category development can be found in Table 1. In the following, the individual categories and the corresponding characteristics will be presented briefly.

The first two categories (Mandate and Sector) are intended to outline the working environment of the organizations. The term "mandate" is used in a variety of ways in the literature, some of which are inconsistent [35]. In the context of this work, the humanitarian or development-oriented nature of the humanitarian organizations' work is referred to. The category Sector describes the main orientation with regard to the objectives of the organizations. It was oriented towards the areas of the cluster approach [36], but also towards specific target groups, such as children. The next two categories describe the size of the organizations, in terms of regional scope (Area) and the available funding

[1] The full list of these organizations can be found in the appendix.

Table 1. Categories for analysis

Category	Attributes
Mandate	(1) humanitarian mandate (2) development mandate
	(3) multi-mandated
Sector	(1) specific sector (2) multi-sectoral
Area	(1) regional focus (2) international
Budget [€]	(1) >250 Mio (2) <250 Mio (3) <100 Mio (4) <50 Mio
	(5) <10 Mio (6) <1 Mio
Public Fund	(1) very low <5% (2) less than 1/3 (3) less than 2/3
	(4) more than 2/3
Network	(1) independent org. (2) int. network (3) alliance organization
	(4) nat. fundraiser
Cooperation	(1) project partner (2) own projects (3) mixed approach
Logistics	(1) no mention (2) particular relevance

(Budget). The regional focus of an organization can be on certain countries, but also on particular regions or continents. If such a focus is not recognizable in the work of an organization, it is classified as international. The attributes according to the available budget are based on the classification according to Harvey et al. [23], but are extended by one group (< €1 million). The classification was carried out, if possible, based on the total budget in 2017[2] and refers in each case to the German humanitarian organization (even if these act within networks). The following two categories are intended to show the dependencies of the organizations - both in financial (Public Fund) and organizational terms (Network). Public funds can be an important source of income for humanitarian actors, but they also imply certain restrictions, as special conditions are often attached [37]. The category of the network is an important distinguishing feature, especially regarding logistical issues. Reference is made to organizational linkage, not to strategic cooperation with other actors. A distinction is made between independent organizations, those operating within an international network, alliances, and national fundraising organizations[3]. Finally, the two categories of the actual working procedures (Cooperation and Logistics) will be considered more in detail. How an organization operates on site has a great influence on the processes and thus also on logistics itself. A distinction is made between organizations that work exclusively with local partners, those that operate autonomously on site (e.g. through their own staff, regional offices, or in a network), and those that do have a mixed approach. Finally, the extent to which logistics is important for the organizations is examined. The focus lies

[2] If the data from 2017 were not available at the time of this study, the corresponding data from previous years were used, which happened nine times.
[3] A detailed description of the individual structures is given in Sect. 4.2.

on whether logistics personnel are employed, own warehouses are operated, or transports from Germany take place.

4 Analysis

4.1 Overview of Humanitarian Organizations in Germany

According to the database of the DZI, the 75 examined German humanitarian organizations are all active in the field of disaster relief[4]. With regard to the analysis criterion of the mandate, there is a clear tendency - in accordance with the literature [35] - towards a multi-mandated approach. A total of 51 of the organizations surveyed see themselves as active in both short-term disaster relief and development cooperation. Almost a quarter of the organizations pursue mainly long-term development cooperation purposes. However, these are also active in exceptional cases in acute disaster operations. Both for development mandated and in some cases for multi-mandated organizations, the focus of work is less on the acute phase immediately after the outbreak of a disaster than on the rehabilitation and reconstruction phase. The opposite is true for organizations with a humanitarian mandate, which are all involved in the acute phase of disaster relief. A total of 8 of the organizations were classified in this category. Slim and Bradley state that very few organizations have a purely humanitarian mandate (ICRC and MSF only) [35], but the study identified more organizations in this category, which may be due to a more comprehensive definition of the term humanitarian. However, the study also shows that organizations with a purely humanitarian mandate form the smallest group (Fig. 1a).

A closer look at the organizations' fields of activity reveals that almost half of the organizations (36) have a multi-sectoral approach. These organizations are engaged in various sectors such as nutrition, shelter, or Water, Sanitation, and Hygiene (WASH) and do not have a specific focus group for their work. The target group supported by the largest share of organizations surveyed (19) are children, often in combination with the sector activity in education. The second largest group (12) is that of organizations working in the health sector. The remaining organizations (11%) each have a unique focus (Fig. 1b). There are organizations which are specified on a particular sector such as WASH (e.g. arche noVa) or Food Security (e.g. Welthungerhilfe) or which support a specific target group such as elderly (e.g. HelpAge) or disabled people (e.g. Handicap International). A recognizable correlation between the mandate and the field of activity of an organization can only be seen in the children/education sector. Organizations active here always include long-term development cooperation projects. In the other fields of activity there is no such clear tendency.

With regard to the regional activities of the organizations, it is evident that a large proportion are internationally active (58). In the case of organizations with

[4] In addition, the organisations are also assigned to other areas of activity, such as development cooperation, human rights or nature, and environmental protection [30].

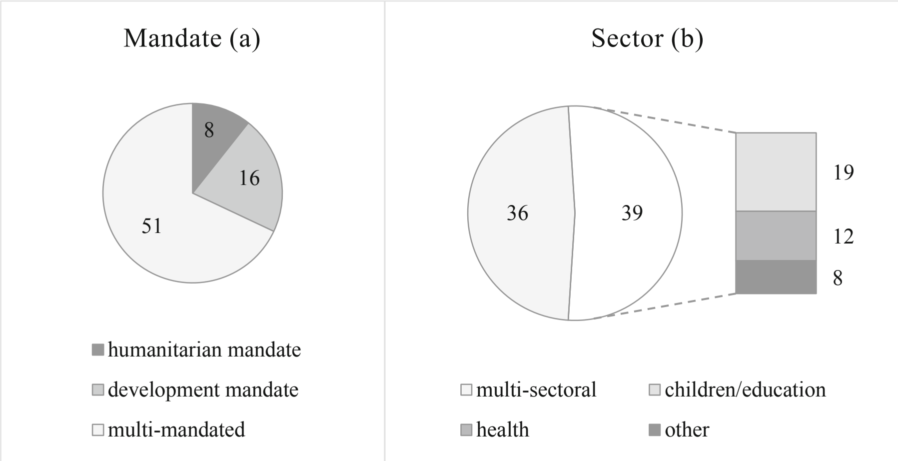


Fig. 1. Analysis of the criteria Mandate and Sector

a regional focus, there are those that work in one or a few specific countries (e.g. Andheri Hilfe in Bangladesh and India) or in a specific region (e.g. Mission East in the Middle East and Asia). A connection with the previous two criteria of analysis can be found with the mandate - all organizations with a humanitarian mandate have an international focus.

The size of the organizations in terms of their budget reveals significant diversity amongst the organizations. The largest organization investigated (Bread for the World), with an annual budget of above €280 million in 2017, is more than 1000 times larger than the smallest organizations (e.g. Helping Hands). With a number of 24 and 28, respectively, organizations with budgets of up to €10 million and up to €50 million account for the largest share. Organizations with an annual budget of more than €50 million represent 17% (13) of the organizations surveyed, but about 75% of the total budget. The corresponding organizations in this group are also exclusively active internationally. Small organizations with a budget of less than €1 million represent 10 (Fig. 2a). Looking at the share of public funds, it is noticeable that almost half (35) receive very little in the way of government subsidies. Some of these organizations also explicitly point out that they operate entirely on the basis of private donations (e.g. MSF). The remaining three groups make up about an equal number with 12 to 15 (Fig. 2b). A correlation between the size of an organization's budget and the share of public funds could not be found. Among the organizations with a very high share of public funds, there are both large (e.g. Welthungerhilfe) and smaller (e.g. NETZ) organizations. The same could be shown for organizations with a very small share of public funds.

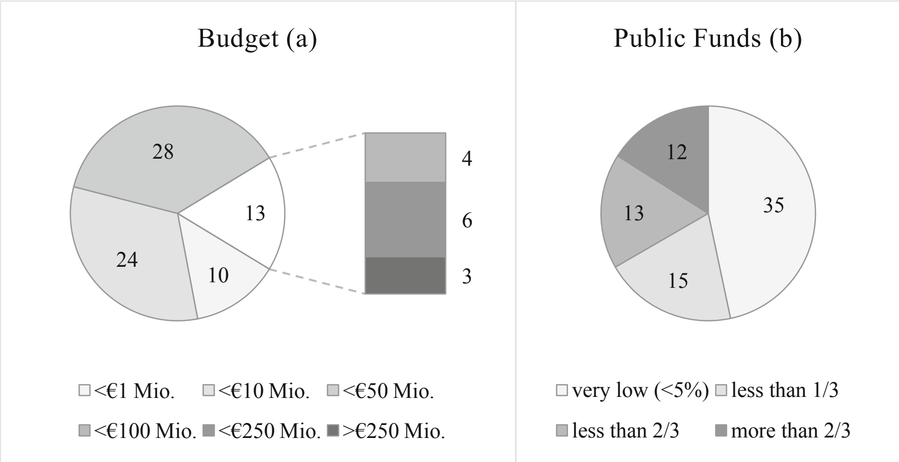


Fig. 2. Analysis of the criteria Budget and Public Funds

4.2 International Networks and Cooperation

There are different factors influencing the processes and consequently the corresponding logistical activities of humanitarian organizations. In the context of this study, the underlying network structure of internationally active organizations was identified as an essential criterion. Six main organizational structures (two of them with two substructures each) were identified, which will be briefly described below (Fig. 3). It should be noted that not the forms of cooperation between different organizations but rather the internal structure of an organization or an organizational network is described. An essential group are those organizations with an international network, where the respective national organizations have independent project responsibility and an umbrella organization does exist, which can either be based in Germany (a) or in another country (b). Examples include Amref Health Africa, Malteser International, or World Vision. Certain substructures can also be formed, as is the case for MSF. Here, individual country sections work more closely together in so-called operational centers. On the other hand, there are international organizational networks in which there is also an organizational head with different country representations, but these do not have their own project responsibility or competences for implementation, but only for financing and informational work (c, d). Examples are the German Committee for UNICEF or Oxfam. Again, there are the two possibilities: the headquarters can be based either in Germany (c) or in another country (d). Furthermore, there are structures in which the organizations work independently in various countries, but as so-called sister organizations there is a certain level of coordination that takes place (e). One example of this is Kindernothilfe. Groups (f) and (g) play a particular role in the network structures. On the one hand, there are network organizations which bring together different national organizations (f), for example to carry out joint fundraising, or to work together in relief operations, such as the German Relief Coalition (Aktion Deutschland Hilft). On the other hand, certain organizations only serve to acquire donations, which in turn are used to support projects of various other humanitarian organizations (g), such as Stiftung RTL. Finally, the group of organizations that act as individual agents is important (h). Even if those organizations do not have any direct internal network partners, cooperation with other organizations is usually even more prevalent. Such connections can be ideological (e.g. in the case of Bread for the World) or contentrelated (e.g. arche noVa). Almost half of the organizations examined are affiliated via the VENRO association mentioned in Sect. 2.1. In addition, there are other alliances, such as an association of organizations in the Catholic Working Group for Emergency and Disaster Relief (KANK). Of the 75 organizations examined, about half are part of internal, international networks, of which more than half have an umbrella organization (a, b). 12 of the organizations are part of an international fundraising association (c, d) and 6 are from network associations with sister organizations (e). The largest group (33) are the independent organizations (h). Alliances organizations (f) and national fundraisers (g) each account for 2 representatives.

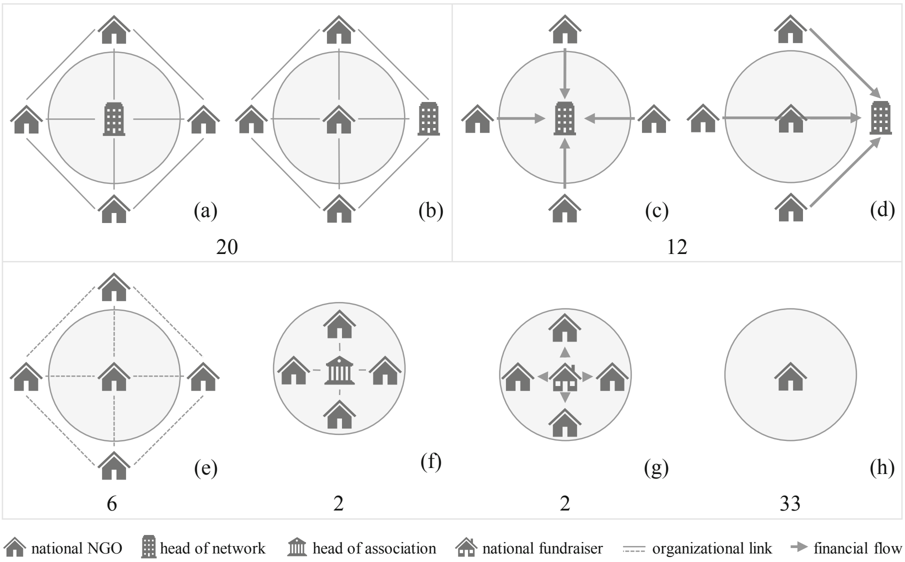


Fig. 3. Characteristics of network structures

A second factor which influences the logistical activities focusses on implementation projects in consideration of the form of cooperation with local partners. Basically, two contradictory directions can be distinguished here: On the one hand, projects can be implemented entirely by local partner organizations or, on the other hand, organizations can implement projects independently on site and with their own staff. In connection with the network structures, this results in different forms. Figure 4 shows the frequencies of the various forms of cooperation that occurred in the context of this study. 36 of the organizations work exclusively through partner organizations, either directly (25) or within their network (9). The remaining 2 imply the two national fundraisers, which exclusively donate funds to other humanitarian organizations. 22 of the organizations surveyed works independently on site, most of them (13) within their own networks. 4 of the organizations worked locally through their umbrella organizations and 5 independently. The remaining 17 of the actors surveyed had a mixed approach of independent work in the countries of deployment and project-related cooperation with partners. Overall, however, only a tendency can be illustrated here with regard to the form of cooperation. Since not all the necessary information could be obtained from the publicly accessible materials, this area requires further research.

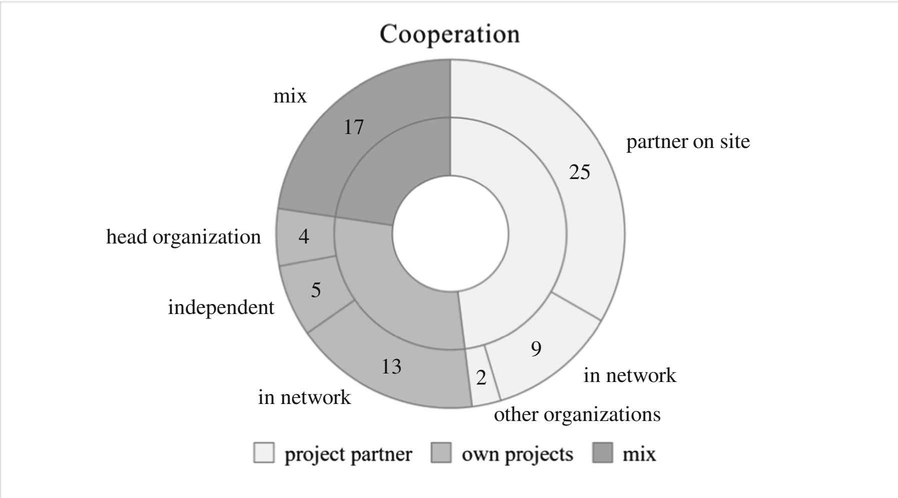


Fig. 4. Analysis of the criteria Cooperation

4.3 Systematization and Logistical Insights

In the final step of the evaluation, conclusions are drawn regarding the relationship between the characteristics of a humanitarian organization and their logistics activities in order to identify classes for systematization. For this purpose, the publicly accessible materials of the 75 organizations were examined with regard to logistics-relevant content. A total of 15 organizations mentioned corresponding information relating to logistical issues. Frequently mentioned in the annual reports (10 times) were logistics personnel or corresponding departments (e.g. ADRA, arche noVa, Welthungerhilfe, and Malteser International). In addition, the same number of reports mentioned aid transports from Germany to the countries of deployment or own storage capacities (e.g. MSF, Aktion Deutschland Hilft, Humedica, and Ora Kinderhilfe). Some organizations refer explicitly to local procurement stragegies or to cooperation with logistics service providers. These tasks are well suited to be assigned to the three main process stages of Balcik et al. [19]. However, the individual passages in the reports are often either very general or refer to specific individual cases. A detailed evaluation is therefore not straightforward. Nevertheless, the relationships between the other organizational characteristics were examined and will be briefly explained in the following.

The mandate of an organization has a major influence on the working methods of an organization. Speed and flexibility are of great importance in relief operations, especially in the acute phase after the outbreak of a disaster, with the result that corresponding logistical challenges increase. A total of 4 of the 15 organizations have a purely humanitarian and only one a development mandate, which differs markedly from the total sample of organizations (cf. Fig. 1a). It could be discovered that actors with a pure development mandate are less interested in logistics with regard to the public communication than those with a humanitarian mandate. It is also important in which disaster management phase an organization is deployed in the event of a disaster. Unfortunately, this information could not be gathered in the context of this study. A further effect on logistical activities seems to be the area of activity of a humanitarian organization. It becomes apparent that the organizations which are active in the health sector place a clear emphasis on logistics (6 organizations). This can be explained by the increased demands in this area. For example, drugs and medical materials are often procured internationally and have special storage requirements with regard to shelf life or temperature. Organizations specialized in helping children, on the other hand, hardly consider logistics to be important (1 organization). With regard to the internal network structure of organisations, a clear trend can be seen in the organizations surveyed. For fundraisers (both within their own networks and those who provide support to various other organizations), logistics plays no role at all, since these actors are not active directly on site. There is no clear tendency for the remaining structural types. However, it can be assumed that this has an influence at least with regard to certain logistical problems, such as the location of warehouses or procurement strategies. Finally, the form of cooperation with local partners is important for the logistical activities. Those organizations working via local project partners on site seem to have significantly fewer logistical challenges than those who work locally themselves. It can therefore be summarized that humanitarian actors can be systematized on the basis of the following four dimensions:

- **Mandate**: in terms of the main focus of an organization and the activity in particular disaster management phases
- **Sector**: in terms of the area of activities of an organization
- **Network**: in terms of the internal structure of organizations within their network

- **Cooperation**: in terms of the autonomy of projects and the cooperation with local partners

Table 2 shows an extract from the complete list of the organizations examined, in order to present an illustrative classification of the humanitarian organizations. The organizations differ in their characteristics in the four dimensions mentioned, whereby there are no two organizations that are identical. However, certain classes can be identified with respect to the dimensions. From a scientific point of view, it seems to be expedient to test existing solutions for applicability within such classes. This could also help to reduce the gap between science and practice, as it would make it easier to develop solutions for a specific, self-consistent groups of humanitarian actors. Furthermore, from a practical viewpoint, organizations which are somewhat similar could be brought together in order to exchange ideas and promote common developments.

Table 2. Excerpt from resulting systematization

Organization	Mandate	Sector	Network	Cooperation
Aktion Deut. Hilft	human.	multi-s.	alliance org.	own proj./net.
Amref Health Africa	develop.	health	int. net./umb. org.	own proj./net.
arche noVa	multi-mand.	WASH	independent	mixed approach
Ärzte ohne Grenzen	human.	health	int. net./umb. org.	own proj./net.
Cap Anamur	multi-mand.	health	independent	own proj./ind.
German Doctors	multi-mand.	health	int. net./sis. org.	mixed approach
Humedica	multi-mand.	health	int. net./umb. org.	own proj./net.
Sternstunden	multi-mand.	multi-s.	nat. fundraiser	proj. part./oth.
UNICEF	multi-mand.	children	int. net./fundr.	own proj./head

5 Conclusion

The purpose of this work was to provide an overview of humanitarian NGOs in Germany which includes criteria relevant to their logistical activities. Furthermore, a classification scheme was identified, enabling the systematization of these organizations. A total of 75 German humanitarian organizations were examined. Publicly accessible data of these organizations have been used for the analysis. In accordance with the theoretical foundations, research criteria were developed. An overview of the actors was followed by a deeper evaluation of the internal network structures and the implementation of projects. In addition, an evaluation of the logistics activities was carried out and a systematization scheme drawn up. As a result of the investigations, it became apparent that there is a wide diversity of different humanitraian organizations, corresponding to a wide variety of aspects determining logistical issues, such as the general orientation of an organizations' work, their size as well as the budget or the international structure of the related network (RQ2). The four dimensions of the mandate and the sector, as well as the internal network structure and the forms of cooperation on site were identified as main influencing factors and can therefore be used to classify humanitarian actors systematically (RQ1).

The results of this work are only based on the publicly accessible information of the humanitarian organizations investigated. Therefore, an exact classification of the individual actors was not always possible. In addition, the classification of NGOs into the corresponding categories is, to a certain extent, based on subjective assessments, here additional investigators would increase validity of the results. Further investigations are necessary in this respect and may include qualitative or quantitative methods. For example, it is necessary to characterize the individual dimensions more distinctly. By means of a qualitative case study with expert interviews, the influences of individual characteristics, such as the network structure, on logistics activities could be revealed. The advantage here would be the use of internal knowledge and the concentration of specific aid organizations. This work can therefore serve as a starting point.

Appendix

Full list of organizations analyzed: ADRA Deutschland e.V.; Ärzte der Welt e.V.; Ärzte ohne Grenzen e.V. - Médecins Sans Frontières (MSF), Deutsche Sektion; Afghanischer Frauenverein e.V.; Aktion Deutschland Hilft e.V.; Aktion Friedensdorf e.V.; Aktion Kleiner Prinz – Internationale Hilfe für Kinder in Not – e.V.; AMREF Deutschland, Gesellschaft für Medizin und Forschung in Afrika e.V.; Andheri – Kinder- und Leprahilfe e.V.; Andheri Hilfe e.V.; arche noVa – Initiative für Menschen in Not e.V.; Bischöfliches Hilfswerk MISEREOR e.V.; Brot für die Welt; Bündnis Entwicklung Hilft – Gemeinsam für Menschen in Not e.V.; Cap Anamur/Deutsche Not-Ärzte e.V.; Christoffel-Blindenmission Deutschland e.V.; CVJM-Gesamtverband in Deutschland e.V. (Aktion Hoffnungszeichen); DESWOS Deutsche Entwicklungshilfe für soziales Wohnungs- und Siedlungswesen e.V.; Deutsche Welthungerhilfe e.V.; Deutscher Caritasverband e.V.; Deutsches Komitee für UNICEF e.V.; Deutsches Rotes Kreuz e.V.; Diakonie Katastrophenhilfe; Difäm – Arbeitszweig Missionsärztliche Dienste; Don Bosco Mission Bonn; Don Bosco Mondo e.V. Jugend.Hilfe.Weltweit; Eritrea-Hilfswerk in Deutschland (EHD) e.V.; Flug-hafenverein München e.V.; German Doctors e.V.; Handicap International e.V.; Help – Hilfe zur Selbsthilfe e.V.; HelpAge Deutschland e.V.; Helping Hands e.V.; Hilfswerk der deutschen Lions e.V.; Human Help Network e.V.; humedica e.V.; Johanniter-Unfall-Hilfe e.V.; Kinderhilfswerk Stiftung Global-Care; Kindermissionswerk "Die Sternsinger" e.V.; Kindernothilfe e.V.; Kinderwerk Lima e.V.; Kirche in Not/Ostpriesterhilfe Deutschland e.V.; LandsAid e.V. – Verein für Internationale Humanitäre Hilfe; Lichtbrücke e.V.; Malteser Stiftung; medico international e.V.; MISSIO – Internationales Katholisches Missionswerk Ludwig Missionsverein KdöR; Mission East Deutschland e.V.; Missionszentrale der Franziskaner e.V.; NETZ Partnerschaft für Entwicklung und Gerech-tigkeit e.V.; Neuapostolische Kirche – karitativ e.V.; nph Kinderhilfe Lateinamerika e.V.; ora Kinderhilfe International e.V.; Oxfam Deutschland e.V.; Plan International Deutschland e.V.; Renovabis e.V.; Samaritan's Purse e.V.; Save the Children Deutschland e.V.; Schülerhilfe für

Nepal e.V.; Shelter Now Germany e.V.; Solidaritätsdienst International e.V. (SODI); SOS-Kinderdörfer weltweit Hermann-Gmeiner-Fonds Deutschland e.V.; Sternstunden e.V.; Stiftung Kinderzukunft; Stiftung Menschen für Menschen – Karlheinz Böhms Äthiopienhilfe; Stiftung RTL – Wir helfen Kindern e.V.; Terra Tech Förderprojekte e.V.; terre des hommes Deutschland e.V. Hilfe für Kinder in Not; Tierärzte ohne Grenzen e.V.; UNO-Flüchtlingshilfe e.V.; Verein zur Förderung der Städtepartnerschaft Kreuzberg – San Rafael del Sur e.V.; Vereinte Evangelische Mission. Gemeinschaft von Kirchen in drei Erdteilen (VEM); World Relief Deutschland e.V.; World Vision Deutschland e.V.; Wort & Tat, Allgemeine Missions-Gesellschaft e.V.

References

1. Kunz, N., van Wassenhove, L.N., Besiou, M., Hambye, C., Kovács, G.: Relevance of humanitarian logistics research: best practices and way forward. Int. J. Oper. Prod. Manag. **37**(11), 1585–1599 (2017). https://doi.org/10.1108/IJOPM-04-2016-0202
2. UNHCR: 3RP Regional strategic overview 2019-20. UNHCR, Geneva (2018)
3. Wallemacq, P., Below, R., McClean, D.: Economic Losses, Poverty & Disasters: 1998–2017. CRED, Brussels (2018)
4. UNISDR: Making development sustainable: The future of disaster risk management. Global assessment report on disaster risk reduction. UNISDR, Geneva (2015)
5. ReliefWeb: Indonesia. https://reliefweb.int/country/idn. Accessed 6 Feb 2019
6. Dijkzeul, D., Reinhardt, D.: Das internationale humanitäre System und seine Akteure. In: Lieser, J., Dijkzeul, D. (eds.) Handbuch Humanitäre Hilfe. Springer, Heidelberg (2013)
7. Chakravarty, A.K.: Humanitarian relief chain: rapid response under uncertainty. Int. J. Prod. Econ. **151**, 146–157 (2014). https://doi.org/10.1016/j.ijpe.2013.10.007
8. Kovács, G., Spens, K.M.: Humanitarian logistics in disaster relief operations. Int. J. Phys. Distrib. Logist. Manag. **37**(2), 99–114 (2007). https://doi.org/10.1108/09600030710734820
9. Baumgarten, H.: Humanitäre Logistik. In: Baumgarten, H., Schwarz, J.R., Keßler, M. (eds.) Humanitäre Logistik: Herausforderungen und Potenziale der Logistik in der humanitären Hilfe. DVV Media Group, Hamburg (2011)
10. Leiras, A., de Brito, I. Jr., Queiroz Peres, E., Rejane Bertazzo, T., Yoshida Yoshizaki, H.T.: Literature review of humanitarian logistics research: Trends and challenges. J. Humanitarian Logist Supply Chain Manag. **4**(1), 95–130 (2014). https://doi.org/10.1108/JHLSCM-04-2012-0008
11. Pedraza-Martinez, A.J., Stapleton, O., van Wassenhove, L.N.: On the use of evidence in humanitarian logistics research. Disasters **37**(1), 51–67 (2013). https://doi.org/10.1111/disa.12012
12. Eberwein, W.-D., Reinalda, B.: A brief history of humanitarian actors and principles. In: Sezgin, Z., Dijkzeul, D. (eds.) The New Humanitarians in International Practice: Emerging Actors and Contested Principles. Routledge, New York (2016)
13. Kovács, G., Spens, K.M.: Trends and developments in humanitarian logistics: A gap analysis. Int. J. Phys. Distrib. Logist. Manag. **41**(1), 32–45 (2011). https://doi.org/10.1108/09600031111101411

14. van Wassenhove, L.N.: Humanitarian aid logistics: Supply chain management in high gear. J. Oper. Res. Soc. **57**(5), 475–489 (2006). https://doi.org/10.1057/palgrave.jors.2602125
15. Thomas, A., Kopczak, L.R.: From logistics to supply chain management: The path forward in the humanitarian sector. (2005). http://www.fritzinstitute.org/pdfs/whitepaper/fromlogisticsto.pdf. Accessed 23 Mar 2018
16. Schwarz, J.R.: Humanitäre Logistik für die Versorgungsproblematik in Subsahara-Afrika: Aufbau von Logistikkapazität durch Wissenstransfer. Dissertation, Technische Universität Berlin (2012)
17. Pettit, S., Beresford, A.: Critical success factors in the context of humanitarian aid supply chains. Int. J. Phys. Distrib. Logist. Manag. **39**(6), 450–468 (2009). https://doi.org/10.1108/09600030910985811
18. Beamon, B.M., Balcik, B.: Performance measurement in humanitarian relief chains. Int. J. Publ. Sect. Manag. **21**(1), 4–25 (2008). https://doi.org/10.1108/09513550810846087
19. Balcik, B., Beamon, B.M., Krejci, C.C., Muramatsu, K.M., Ramirez, M.: Coordination in humanitarian relief chains: Practices, challenges and opportunities. Int. J. Prod. Econ. **126**(1), 22–34 (2010). https://doi.org/10.1016/j.ijpe.2009.09.008
20. Hellingrath, B., Widera, A.: Prozessgestaltung in der humanitären Logistik: Bericht der AG Prozesse. In: Baumgarten, H., Schwarz, J.R., Keßler, M. (eds.) Humanitäre Logistik: Herausforderungen und Potenziale der Logistik in der humanitären Hilfe. DVV Media Group, Hamburg (2011)
21. Deloffre, M.Z.: Global accountability communities: NGO self-regulation in the humanitarian sector. Rev. Int. Stud. **42**(4), 724–747 (2016). https://doi.org/10.1017/S0260210515000601
22. Development Initiatives: Global Humanitarian Assistance Report 2018. Development Initiatives, Bristol (2018)
23. Harvey, P., Stoddard, A., Harmer, A., Taylor, G.: The State of the Humanitarian System: Assessing Performance and Progress. ALNAP/ODI, London (2010)
24. IFRC: The movement. http://www.ifrc.org/en/who-we-are/the-movement/. Accessed 8 Feb 2019
25. ALNAP: The State of the Humanitarian System. ALNAP/ODI, London (2018)
26. Weingärtner, L., Otto, R.: Die deutsche humanitäre Hilfe. In: Lieser, J., Dijkzeul, D. (eds.) Handbuch Humanitäre Hilfe. Springer, Heidelberg (2013)
27. Deutscher Bundestag: Bericht der Bundesregierung über die deutsche humanitäre Hilfe im Ausland 2014 bis 2017. Drucksache 19/5720 (2018)
28. VENRO: VENRO-Arbeitsgruppen. https://venro.org/ueber-venro/arbeitsgruppen/. Accessed 8 Feb 2019
29. Eisenhardt, K.M.: Building theories from case study research. Acad. Manag. Rev. **14**(4), 532–550 (1989). https://doi.org/10.5465/amr.1989.4308385
30. DZI: Was kann die Spenderberatung für Sie tun? https://www.dzi.de/spenderberatung/. Accessed 7 Jan 2019
31. Yin, R.K.: Case Study Research: Design and Methods, 5th edn. Sage, Los Angeles (2013)
32. Stuart, I., Mc Cutcheon, D., Handfield, R., McLachlin, R., Samson, D.: Effective case research in operations management: A process perspective. J. Oper. Manag. **20**(5), 419–433 (2002). https://doi.org/10.1016/S0272-6963(02)00022-0
33. Ellram, L.M.: The use of the case study method in logistics research. J. Bus. Logist. **17**(2), 93–138 (1996)
34. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken, 12th edn. Beltz, Weinheim (2015)

35. Slim, H., Bradley, M.: Principled Humanitarian Action & Ethical Tension in Multi-mandete Organizations. World Vision International, Geneva (2013)
36. IASC: Reference module for cluster coordination at country level. IASC, Geneva (2015)
37. Schweizer, B.: Humanitäre Dilemmata: Anspruch und Wirklichkeit der humanitären Prinzipien. In: Lieser, J., Dijkzeul. D. (eds.) Handbuch Humanitäre Hilfe. Springer, Heidelberg (2013)

Supply Chain Operations

An Integrated Multi-criteria Approach for the Regional Facility Location and Development Planning Problem

David Kik(✉), Matthias G. Wichmann, Christoph Johannes, and Thomas S. Spengler

Technische Universität Braunschweig, Brunswick 38106, Germany
d.kik@tu-braunschweig.de

Abstract. Urbanization trends confront companies with major challenges in their location management. Facility location planning is a complex task, as it usually involves multiple conflicting objectives. In addition, particularly at a regional decision-making level, companies are confronted with relevant dynamics that cause main changes of location requirements and characteristics over time. In order to ensure optimal long-term location decisions, planning approaches have to consider company-driven and municipal location developments over an extended planning horizon. To date, the variety of dynamic multi-criteria facility location planning approaches is generally scarce and possibilities for location developments are not given. Therefore, an appropriate model formulation for regional facility location planning considering relevant dynamics in location development is missing. In this contribution, the Regional Facility Location and Development Planning Problem (RFLDP) is introduced. For this purpose, an appropriate model is developed that provides an integrated decision-aid for a company's location selection and development. In doing so, a strategic measure plan is determined through an allocation of company-driven measures under consideration of municipal location developments. The model is evaluated based on an illustrative example.

Keywords: Facility location planning · Measure allocation · Municipal developments

1 Introduction

Ongoing urbanization is characterized by trends significantly influencing the characteristics of existing cities such as expandation to so-called megacities, rising migrations and thus increasing urban densities. As a result, the interweaving between urban and rural areas grows continuously, the urban production expands and demand on resources in surrounding areas increases (United Nations 2014). Therefore, companies producing on the outskirts of cities are confronted with main challenges in location management, as locations continue to develop and differ significantly in their characteristics within a region.

Therefore, companies are faced with dynamic changes over time. These dynamics are threefold. First, the company's location requirements change over time. This is

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 241–254, 2019.
https://doi.org/10.1007/978-3-030-29821-0_16

usually due to technological improvements, e.g. the evolution of broadband speed over time. Second, characteristics of regional location factors can be expected to change over time due to location developments. Here, location factors are mainly influenced and actively changed due to municipal location developments as well as company-driven development measures. Third, interdependencies between development measures are possible. For example, in a joint broadband infrastructure development project, a company could benefit from synergetic interdependencies such as lower fixed costs through sharing excavation costs with the municipality. Thus, a company could invest more money in a broadband infrastructure's performance.

Therefore, many regional location decisions require a close interlinking between municipalities and companies who should see themselves as cooperative partners in regional location planning and development.

So far, however, the focus in facility location planning is the choice of a suitable location that best meets the company's location requirements, rather than on the development of locations in a multi-criteria environment. Therefore, there is no appropriate model formulation that provides a target-oriented development of locations considering relevant dynamics and ensures a long-term optimal decision in regional facility location planning.

In this paper we present a multi-criteria model for a regional facility location and development planning problem (RFLDP). The novel planning approach allows for an integrated location selection and allocation of company-driven measures, in order to improve locations according to a company's requirements over time. In addition, municipal decisions in location development are taken into account.

The paper is structured as follows: Sect. 2 addresses the main problem characteristics and gives a brief literature overview. In Sect. 3 a model formulation for the RFLDP is introduced. In Sect. 4 an illustrative example is given, in order to highlight the potentials of the novel approach. Finally, the paper closes with a conclusion and an outlook.

2 Regional Facility Location and Development Planning Problem (RFLDP)

2.1 Problem Characteristics

The aim in facility location planning is to determine suitable locations that meets a company's requirements in the best possible way under consideration of resource restrictions. A well-located location is essential for a company's sustainable success. A location decision involves high overall risks and requires an accurate planning, as it is of strategic and long-term nature. The decision is also complex, since it is a multi-criteria problem with usually conflicting objectives (Current et al. 1990). The importance of individual location factors differs between decision-making levels. While on a global level monetary factors (i.e. labor costs) are dominating, on a regional level non-monetary factors (i.e. staff qualification) are becoming increasingly relevant (Abele et al. 2006).

Moreover, due to ongoing urbanization trends, companies are faced with relevant dynamics in regional location planning and development. These dynamics are threefold. First, a company's location requirements usually change over time. This is usually due to economical and technological improvements. Second, location factor characteristics can also be expected to change over time. This has two reasons. On the one hand, location factor characteristics are mainly influenced by municipal location developments. On the other hand, a company has opportunities to improve location factors actively by allocating suitable measures in a target-oriented manner. Third, there may be possible interdependencies between measures while developing locations. In this context, in most recent empirical studies synergetic and conflicting interdependencies are identified (Krause et al. 2018, Kik et al. 2018a).

In particular, an appropriate planning model for the RFLDP problem must provide the designing of locations, in order to select a suitable location and develop it according to a company's requirements over an extended planning horizon considering multiple criteria.

2.2 Literature Overview

In location planning, companies deal with decision-making situations such as selecting and designing suitable facility locations. According to Current et al. (1990), facility location planning models can be characterized by five structural aspects. First, regarding the number of facilities sought (e.g. single- or multi-facility location problem). Second, whether the facilities to be located are in any way capacitated or not (e.g. production capacities). Third, is the number of potential facilities pre-determined or not (e.g. discrete or continuous decision space). Fourth, are the given parameters deterministic or stochastic (e.g. certain or uncertain demands). Finally, considering a single period with constant parameters or multiple periods with over time changing parameters (e.g. static or dynamic).

Since, location planning is usually a problem with multiple and often conflicting criteria, multi-criteria decision making (MCDM) approaches seems to be appropriate for that purpose. Detailed reviews of MCDM applications in facility location planning for example are given by Farahani et al. (2010), Hekmatfar and SteadieSeifi (2009) and Current et al. (1990). Multi-criteria facility location planning approaches can be classified in two streams. The first stream includes multi-objective (MODM) approaches, aiming to design the best alternative considering quantifiable objectives and well-defined constraints. The second stream includes multi-attribute (MADM) approaches, trying to select the most suitable alternative considering more or less imprecise attributes. For an overview of MADM applications in facility location planning, see the above-mentioned review of Farahani et al. (2010).

Since, location developments over time are explicitly considered in RFLDP and therefore a design problem of dynamic nature is implied, some existing dynamic approaches of the former stream are depicted below. Xu et al. (2016) introduce a multi-objective dynamic model for the location planning of hydropower stations in a discrete and fuzzy environment. Dias et al. (2008) propose a multi-objective dynamic approach to open, close and reopen facilities considering an allocation of clients to facilities over an extended planning horizon. Melachrinoudis and Min (2000) propose a multi-objective formulation for a dynamic facility relocation problem incorporating the transfer of production capacities from existing facilities to new facilities over time. In earlier research also Min and Melachrinoudis (1996) provide a dynamic goal programming approach for a global location planning problem incorporating multiple periods, over the time changing as well as stochastic factors. Reeves et al. (1988) applied a dynamic goal programming model to an industrial capacity expansion problem involving existing as well as new facilities under consideration of over the time changing demands.

Basically it can be noted that the variety of MODM approaches incorporating dynamics in facility location planning is relatively scarce. Apart that fact, no possibilities for a company's target-oriented influence in location development through an allocation of measures are taken into account in existing facility location models. Given the potentials, it seems inevitable to integrate the allocation of measures in regional location planning (Kik et al. 2018b, Kik et al. 2018a, and Richter 2017). Therefore, following the introduced structural characteristics in facility location planning, there is no appropriate model formulation that meets the requirements of a RFLDP.

3 Model

3.1 Problem Setting, Approach and Assumptions

The RFLDP addressed in this contribution is a facility location problem in a multi-criteria environment. According to the introduced structural characteristics of facility location problems, the RFLDP is a single- (non-capacitated) facility location problem. The decision space has discrete as well as continuous components. On the one hand, there is a discrete decision space in location selection, since the location alternatives are explicitly predetermined. On the other hand, the decision space in allocating company-driven measures is continuous, since the characteristics of the locations to be developed are not predetermined and have to be calculated (see assumption 4 below). Since parameters are assumed to be known and change over an extended planning horizon, the RFLDP is of deterministic and dynamic nature.

The main planning tasks in the RFLDP are an integrated selection of a compromise-optimal facility location and to determine a strategic location development plan by allocating company-driven measures in a target-oriented manner (see Fig. 1).

The integrated model to be developed for the RFLDP problem is based on the weighted goal programming (GP) model formulation. In general, GP is still the most widely used multi-criteria technique with extensive applications in especially management science and engineering (Colapinto 2015, Tamiz et al. 1998). Considering the fact that a simultaneous achievement of maximum values in multiple conflicting criteria problems such as location planning is usually utopian, the GP is a far more appropriate approach comparing to other MODM approaches. That is why GP aiming rather the best possible achievement of defined targets (satisficing principle) than an achievement of maximum values (optimizing principle) (Ottmann and Lifka 2010). In particular, in multi-criteria environments satisficing describes the decision-making behavior far more accurately than rational optimizing. Moreover, compared to other MODM approaches, GP is more simple as well as flexible in modeling and solution finding (Tamiz et al. 1998).

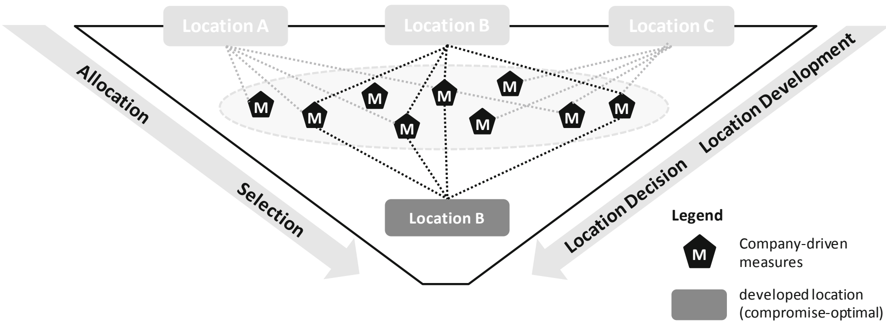


Fig. 1. Integrated planning tasks in the RFLDP

Due to the flexibility of GP, different types of decision-relevant objectives can be applied. Especially, the weighted GP provides more flexibility in order to include decision-maker's preferences and the possibility to perform trade-offs among multiple conflicting factors, which are both desired features in facility location planning.

The integrated planning tasks in RFLDP are based on the following five assumptions:

Assumption 1: Usually strong differing units of qualitative and quantitative regional location factors are normalized, in order to ensure direct comparability among each other.
Assumption 2: The decision-maker is aware of his preferences. Therefore, the weightings of the regional location factors are known, which are absolute values.
Assumption 3: The location decision as well as the location development plan for all planning periods are made in the present planning period t_0.
Assumption 4: It is assumed that development investment for company-driven measures increase proportionally with their growing intensities. Therefore, the above mentioned continuous decision space regarding the allocation of company-driven measures is justified.
Assumption 5: It is assumed that the effectiveness of company-driven measures as well as municipal developments and their effectiveness on location factors are known and therefore deterministic.

3.2 Notation

The following notation is used here to define indices, decision as well as auxiliary variables and parameters for the RFLDP model formulation.

Indices	
i	Index of regional location factors (i.e. goals), $i = 1, \ldots, I$
j	Index of potential locations, $j = 1, \ldots, J$
m	Index of company-driven measures, $m = 1, \ldots, M$
s	Index of municipal developments, $s = 1, \ldots, S$
t, t'	Index of periods, $t, t' = 0, \ldots, T$

Parameters	
$a_{i,j}$	Characteristic of a location factor i at location j in t_0
A_i^{min}	Desired minimum characteristic of a location factor i
$b_{i,t}$	Target for a location factor i in period t
B	Overall available budget for the construction and development of a location
$CE_{i,m}$	Effectiveness of a company-driven measure m on location factor i
d_t	Discount rate in period t, with $d_t = (1+h)^{-t}$
h	Calculatory interest rate
IC_j	Construction investment for a location j
$ID_{m,t}$	Development investment for the allocation of a company-driven measure m in period t
M	Very large number, known as "Big M"
$ME_{i,j,s,t}$	Effectiveness of a municipal development s initiated at site j on location factor i in period t
$w_{i,t}^{+}$	Relative weighting of an undesired positive target deviation for location factor i in period t
$w_{i,t}^{-}$	Relative weighting of an undesired negative target deviation for location factor i in period t
Decision and auxiliary variables	
$d_{i,j,t}^{+}$	Positive deviation from target for location factor i in period t at location j
$d_{i,j,t}^{-}$	Negative deviation from target for location factor i in period t at location j
$D_{i,j,t}$	Buffer variable to compensate difference between actual and targeted characteristics of location factor i at location j in period t
$x_j = \begin{cases} 1, \\ 0, \end{cases}$	if potential location j is selected otherwise
$y_{j,m,t}$	Degree of intensity of an allocated company-driven measure m to potential location j in period t

3.3 Formulation

Having declared the notation, the RFLDP model formulation is introduced. The model formulation is based on the weighted GP formulation by Zimmermann and Gutsche (1991). Therefore, the extended model consists of the adapted objective function (1) and modified constraints (c.f. Formulas (3), (13)) as well as new added constraints (c.f. Formulas (2), (4)–(10), (11)–(12)).

$$Min\,Z = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=0}^{T}\left(w_{i,t}^{+}\,d_{i,j,t}^{+} + w_{i,t}^{-}d_{i,j,t}^{-}\right) \tag{1}$$

The objective function (1) minimizes the weighted sum of undesirable deviations from the decision maker's targets for all decision-relevant normalized location factors i, which are also location- and time-dependent. Therefore, all location factors i are considered simultaneously. A deviation from the target can be either positive $d^{+}_{i,j,t}$ or negative $d^{-}_{i,j,t}$. Depending on the type of a location factor, just one deviation (i.e. one-sided goal) or both deviations (i.e. two-sided goals) are undesirable, which are penalized by the relative weightings $w^{+}_{i,t}$ and $w^{-}_{i,t}$. These represents the decision-maker's time-dependent preferences regarding decision-relevant location factors.

In total, the RFLDP model consists of the following 11 constraints that are structured in five categories of constraints. The first category address main admissibilities when selecting a location j.

$$\sum_{j=1}^{J} x_j = 1 \tag{2}$$

Constraint (2) ensures that only one location is selected by setting the sum of the decision variable x_j to 1. According to the introduced assumptions, the location j is selected in the present planning period t_0.

The second category encompasses the determination of the deviation variables $d^{+}_{i,j,t}$ and $d^{-}_{i,j,t}$ for each location factor i by comparing the achieved and targeted location factor characteristics in period t. The main challenge is a correct calculation of the deviation variables.

$$\begin{aligned} & a_{i,j}x_j + \sum_{t'=0}^{t}\sum_{m=1}^{M} y_{j,m,t'} CE_{i,m} + \sum_{t'=0}^{t}\sum_{s=1}^{S} ME_{i,j,s,t'} && \forall i \in I, \\ & -d^{+}_{i,j,t} + d^{-}_{i,j,t} + D_{i,j,t} = b_{i,t} && j \in J, \\ & && t \in T \end{aligned} \tag{3}$$

Constraint (3) determines the achieved location factor characteristics in period t. These are compared to the targets $b_{i,t}$ for each period. The former is constituted by the given characteristics in t_0 and the changes due to location developments up to the current period t. First, the location factor characteristics in t_0 are calculated by multiplying the decision variable for location selection x_j with the parameter $a_{i,j}$ describing the characteristics of a location factor i at location j. Second, the effectiveness of company-driven measures as well as municipal decisions in location development on the characteristics of location factors are taken into account. On the one hand, the total effectiveness of company-driven measures on a location factor i is calculated through the sum of multiplied decision variables $y_{j,m,t}$ with the parameter for measure effectiveness $CE_{i,m}$. On the other hand, the total effectiveness of municipal developments s on a location factor i is calculated by summing all corresponding parameters for municipal effectiveness $ME_{i,j,s,t}$. In order to incorporate company-driven and municipal developments taken place in previous periods, a sum of the present period t_0 up to a considered period is introduced for both types of location developments.

In order to meet the previously mentioned challenge of correct calculation of the deviation variables $d^{+}_{i,j,t}$ and $d^{-}_{i,j,t}$ that are used by constraint (3), the following value-defining constraints are introduced. Constraints (4) and (5) ensure that all deviation variables can have a value greater than 0 in the case of a selected location j. Otherwise, the deviation variables have a value equal 0.

$$d^{+}_{i,j,t} \leq x_j \cdot M \quad \forall i \in I, j \in J, t \in T \tag{4}$$

$$d^{-}_{i,j,t} \leq x_j \cdot M \quad \forall i \in I, j \in J, t \in T \tag{5}$$

To ensure that constraint (3) remains fulfilled due to these restrictions, a dummy variable $D_{i,j,t}$ is introduced. The dummy variable has a compensating function regarding the difference between the actual and targeted location factor characteristics.

$$D_{i,j,t} \leq \left(-x_j + 1\right) \cdot M \quad \forall i \in I, j \in J, t \in T \tag{6}$$

$$D_{i,j,t} \geq \left(x_j - 1\right) \cdot M \quad \forall i \in I, j \in J, t \in T \tag{7}$$

Therefore, its value depends on the values of the calculated deviation variables. Constraints (6) and (7) ensure that $D_{i,j,t}$ is equal 0 in case of a selected location. Otherwise, its value corresponds to the difference between the actual and targeted location factor characteristics.

The third category address main admissibilities in the allocation of company-driven measures m.

$$y_{j,m,t} \leq x_j \quad \forall j \in J, m \in M, t \in T \tag{8}$$

Constraint (8) ensure that a company-driven measure m is allocated to a location j in period t if a location j is actually chosen.

The fourth category contains essential requirements for the integrated model.

$$a_{i,j}x_j + \sum_{t'=0}^{t}\sum_{m=1}^{M} y_{j,m,t'}CE_{i,m} + \sum_{t'=0}^{t}\sum_{s=1}^{S} ME_{i,j,s,t'} \geq A_i^{min}x_j \qquad \begin{matrix}\forall i \in I,\\ j \in J,\\ t \in T\end{matrix} \tag{9}$$

Constraint (9) ensures that actual location factor characteristics fulfill a desired minimum characteristic A_i^{Min} over all periods. The majority of the left-hand-side formulation is equal to that of constraint (3). By doing so, the parameter A_i^{Min} is multiplied with the decision variables describing the location selection x_j. This ensures a valid solution, although some locations are not chosen.

$$x_jIC_j + \sum_{m=1}^{M}\sum_{t=0}^{T}\left(y_{j,m,t}ID_{m,t}\right) \cdot d_t \leq B \qquad \forall j \in J \tag{10}$$

Constraint (10) ensures that construction investments IC_j of a location j as well as the sum of development investments $ID_{m,t}$ for the allocation of company-driven measures m do not exceed a restrictive overall budget B. In addition, future development investments $ID_{m,t}$ are discounted to the decision time t_0 by a multiplication with corresponding time-dependent discount rates d_t.

Finally, the model is completed by the fifth category including system constraints.

$$x_j \in \{0, 1\} \quad \forall j \in J \tag{11}$$

$$y_{j,m,t} \in [0, 1] \quad \forall j \in J, m \in M, t \in T \tag{12}$$

$$d^+_{i,j,t}, d^-_{i,j,t} \geq 0 \quad \forall i \in I, j \in J, t \in T \tag{13}$$

The constraint (11) describes the binarity of the decision variable for location selection x_j. Constraint (12) sets the lower and upper bound of the continuous decision variable for intensities of company-driven measures $y_{j,m,t}$. Lastly, constraint (13) ensures the non-negativity of both deviation variables $d^+_{i,j,t}$ and $d^-_{i,j,t}$.

All constraints introduced are linear constraints. Therefore, the RFLDP model can be formulated as a mixed integer linear program (MILP).

4 Model

In this section an illustrative example is given, in order to show the potentials of the novel RFLDP model considering a target-oriented allocation of company-driven measures in location development. To achieve numerical results, the RFLDP model is implemented in CPLEX 12.8 and solved on a 2.7 Ghz CPU with 16 GB RAM.

In the example, an automotive supplier seeks a suitable location for its new battery cell production facility in a German region and wants to determine an optimal strategic location development plan, in order to expand its production capacities. There is a total of $J = 3$ potential locations, each of which requires different construction investments (IC_j). Moreover, the locations differ regarding the characteristics of $I = 4$ regional location factors. These are namely the availability of industrial areas ($i = 1$), broadband infrastructure $(i = 2)$, qualification of employees $(i = 3)$ and location attractiveness $(i = 4)$. The location factors are normalized and their characteristics $a_{i,j}$ are rated on a scale from 1 (very poor) to 6 (very good). Since the automotive suppliers' requirements and preferences change over $T = 2$ periods, the time-dependent targets for location factors $b_{i,t}$ and preferences regarding the corresponding deviation variables $\left(w^+_{i,t}, w^-_{i,t}\right)$ are given for both periods. For example, it can be seen for t_0 that the broadband infrastructure is four times more important than the qualification of employees. In t_1 the supplier attaches more importance to a specialized qualification required for the production planning and manufacturing of battery cells. The company's location decision as well as the location development plan over all planning periods are made in the present period t_0. The supplier has a total of 10 million euros at his disposal for the location selection and development. The above mentioned and further relevant parameters are given in Table 1.

Moreover, the automotive supplier considers relevant regional dynamics in location planning that have a significant influence on relevant location factor characteristics.

Table 1. Location factors – characteristics, requirements and preferences

	$a_{i,j}$			$b_{i,t}$		$w^+_{i,t}$	$w^-_{i,t}$	$w^+_{i,t}$	$w^-_{i,t}$
	$j=1$	$j=2$	$j=3$	t_0	t_1	t_0		t_1	
$i=1$	3	3	4	3	4	0	3	0	3
$i=2$	3	5	3	5	6	0	4	0	1
$i=3$	5	4	2	5	6	0	2	0	2
$i=4$	3	2	3	5	5	0	1	0	4
IC_j (in Tsd. Euro)	7900	8400	8200						

On the one hand, $S = 3$ future municipal developments in t_1 at the potential locations are taken into account. These are the expansion and restoration of industrial areas ($s = 1$), support for broadband expansion through a cooperation with a regional energy supplier ($s = 2$), establishment of a pre-school ($s = 3$). The municipal developments $s = 1$ and $s = 3$ are planned at locations $j = 1$ and $j = 2$, while $s = 2$ and $s = 3$ are planned at $j = 3$. On the other hand, in both periods the automotive supplier can take target-oriented influence on location factor characteristics due to $M = 3$ company-driven measures. These are financial support in broadband expansion ($m = 1$), further training of employees ($m = 2$), creation of an attractive living space ($m = 3$). Further, the percentage intensity of each company-driven measure can be determined in each period t. The effectiveness of company-driven measures as well as municipal decisions on location factors are given in Table 2. Further, the corresponding development investments $ID_{m,t}$ (in Tsd. Euro) for company-driven measures for both periods are given in Table 3. Future development investments in period t_1 are given as net values, since they are discounted with an interest rate of 5% per annum to the decision period t_0. In this example the given parameters in Tables 2 and 3 are assumed to be equal for all potential locations j.

Table 2. Effectiveness of location developments on location factor characteristics

	$m=1$	$m=2$	$m=3$	$s=1$	$s=2$	$s=3$
$i=1$	-	-	-	0,7	-	-
$i=2$	1	-	-	-	0,7	-
$i=3$	0,7	0,5	-	-	0,6	-
$i=4$	0,6	-	0,9	-	0,7	0,4

With the given database a problem instance results with 177 variables (3 integer), 271 constraints and 746 nonzeros. To solve this problem, the CPLEX solver needs 0.05 s. Thus, our RFLDP model proves to be an adequate tool for small problem instances.

In order to emphasize the potentials of our developed RFLDP model, the optimal location is determined in two different ways.

Table 3. Development investments for company-driven measures (in Tsd. Euro)

$ID_{m,t}$	$m = 1$	$m = 2$	$m = 3$
t_0	900	600	500
t_1	860	570	480

First, a conventional location decision is made according to existing facility location planning models. It is made based only on period t_0. Thereby, company-driven measures as well as municipal decisions in location development are not taken into account. We call this the initial solution. Second, a location decision is made using our RFLDP model. In contrast, this decision is made based on an extended planning horizon and encompasses all decision-relevant periods. In Particular, it provides a target-oriented allocation of company-driven measures considering municipal decisions in location development.

Since the objective function values are obtained in both differing framework conditions, a direct comparison is not given. Therefore, a comparability is to be ensured through an objective-function-oriented evaluation. In doing so, the initial solution is transformed into the framework conditions of our RFLDP model considering all periods and municipal developments taking place anyway. In this case, only the possibility of allocating company-driven measures is not given. The obtained results are given in Fig. 2.

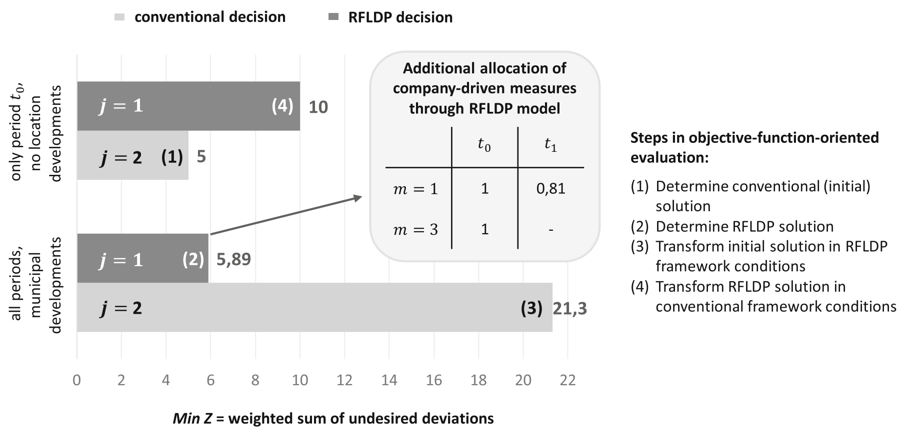


Fig. 2. Comparison of obtained objective function values

In general, in direct comparison it can be noticed that two different location decisions are made. On the one hand, following the conventional location decision, location $j = 2$ (Min Z = 5) is chosen considering only period t_0 without any location developments. As noted above, this is the initial solution (1). On the other hand, our RFLDP model decides for location $j = 1$ (Min Z = 5,89) considering municipal as well as company-driven location developments over an extended planning horizon (2). With the establishment of a strategic action plan and the associated allocation of company-driven, location factor characteristics at location $j = 1$ are improved in a target-oriented manner over time. In doing so, measures $m = 1$ and $m = 3$ are each allocated in full intensity (100%) in t_0, while in t_1 measure $m = 1$ is allocated with a slightly reduced intensity of 81%.

As already mentioned, the two objective function values cannot be directly compared due to different framework conditions in decision-making. For that reason, the initial solution is transformed into the framework conditions of our RFLDP model and a new objective function value is obtained (3). Thereby, an extended planning horizon and municipal location developments are also considered. However, the possibility of allocating company-driven measures is not given in conventional decision-making. Due to this objective-function oriented evaluation a direct comparison is ensured and it can be clarified that the location decision using our RFLDP model obtains a significantly better objective-function value (Min Z = 5,89) than under conventional decision making (Min Z = 21,3). Accordingly, it can be stated that the location chosen in conventional decision-making ($j = 2$) is not the optimal location decision from a long-term perspective. Transforming the RFLDP solution into the framework conditions of conventional decision-making (4), it can be noted that location $j = 1$ (Min Z = 10) performs worse than location $j = 2$ (Min Z = 5) considering only period t_0. However, location $j = 1$ performs way better than $j = 2$ over an extended planning horizon by taking municipal as well as company-driven location developments into account and is optimal from a long-term point of view.

In conclusion, it can be shown that our RFLDP model has major potentials compared to a conventional decision in regional facility location planning. Through the determination of a strategic measure plan locations are developed in a target-oriented manner over time and significantly better location factor characteristics can be achieved. Taking into account time-dependent dynamics in regional location planning for the first time, a significant contribution can be made to an optimal long-term location decision.

5 Conclusion and Outlook

Ongoing urbanization faces companies with major challenges in regional location management and requires a re-thinking in municipal urban development. In order to exploit main potentials of urbanization on both sides, companies and municipalities must henceforth see each other as development partners and cooperate. In this contribution, the Regional Facility Location and Development Planning Problem (RFLDP) is introduced and an appropriate multi-criteria model formulation is presented. The RFLDP model provides an integrated selection of suitable facility locations and the determination of strategic measure plans, due to an allocation of company-driven measures under consideration of municipal decisions in location development. Compared to conventional decision-making in facility location planning, first findings of an illustrative example emphasize major potentials of our RFLDP model. Therefore, the RFLDP model provides significant better location factor characteristics over an extended planning horizon and makes a significant contribution in optimal location decisions from a long-term point of view.

However, further research is conducted on our RFLDP that is threefold. First, the RFLDP model should be validated on the basis of larger problem instances from regional location planning practice. Second, it should be examined how interdependencies between measures for location development in the case of non-additive marginal benefits can be incorporated, if measures influence each other disproportionately regarding their effectiveness. Third, the obtained solutions should be validated for robustness through a scenario oriented-planning based on expressive case studies in regional location planning. Thereby, further research is being conducted on practical model extensions such as a rolling measure planning in order to provide the decision maker more flexibility so that adjustments to the strategic measure plan can be made over time. Furthermore, it is being examined how uncertainties of external parameters can be handled appropriately such as occurrence and effectiveness of municipal location developments.

This work is supported by the German Research Foundation [grant number SP 491/8-1].

References

Abele, E., Kluge, J., Näher, U. (eds.): Handbuch Globale Produktion. Carl Hanser Verlag, München (2006)

Colapinto, C., Jayaraman, R., Marsiglio, S.: Multi-criteria decision analysis with goal programming in engineering, management and social sciences: a state-of-the art review. Ann. Oper. Res. **251**(1–2), 7–40 (2015)

Current, J., Min, H., Schilling, D.: Multiobjective analysis of facility location decisions. Eur. J. Oper. Res. **49**(3), 295–307 (1990)

Dias, J., Captivo, M.E., Climaco, J.: A memetic algorithm for multi-objective dynamic location problems. J. Global Optim. **42**(2), 221–253 (2008)

Farahani, R.Z., SteadiSeifi, M., Asgari, N.: Multiple criteria facility location problems: a survey. Appl. Math. Model. **34**(7), 1689–1709 (2010)

Hekmatfar, M., SteadieSeifi, M.: Multi-criteria location problem. In: Farahani, R.Z., Hekmatfar, M. (eds.) Facility Location: Concepts, Models, Algorithms and Case Studies, pp. 373–393. Physica-Verlag, Heidelberg (2009)

Kik, D., Krause, M., Wichmann, M., Spengler, T., Nyhuis, P.: Betriebliche Standortplanung und -entwicklung in Metropolregionen – Einblicke in die unternehmens- und kommunenseitige Beeinflussbarkeit qualitativer und quantitativer regionaler Standortfaktoren. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, Band, vol. 113(7–8), pp. 503–507. Carl Hanser Verlag, München (2018a)

Kik, D., Wichmann, M.G., Krause, M., Nyhuis, P., Spengler, T.S.: Multikriterielle Entscheidungsunterstützung in der regionalen Standortplanung und -entwicklung. In: OR News, November 2018 (64), pp. 6–11 (2018b)

Krause, M., Kik, D., Nyhuis, P., Spengler, T.S.: Betriebliche Standortplanung und -entwicklung in Metropolregionen–Ergebnisse empirischer Unternehmens- und Kommunenbefragungen, Herausgegeben vom Institut für Fabrikanlagen und Logistik, Institut für Automobilwirtschaft und Industrielle Produktion. Hannover, Braunschweig (2018)

Melachrinoudis, E., Min, H.: The dynamic relocation and phase-out of a hybrid, two-echelon plant/warehousing facility: a multiple objective approach. Eur. J. Oper. Res. **123**(1), 1–15 (2000)

Min, H., Melachrinoudis, E.: Dynamic location and entry mode selection of multinational manufacturing facilities under uncertainty: a chance-constrained goal programming approach. Int. Transp. Oper. Res. **3**(1), 65–76 (1996)

Ottmann, M., Lifka, S.: Methoden der Standortanalyse. In: Haas, H.-D. (ed.) Geowissen kompakt. Wissenschaftliche Buchgesellschaft, München (2010)

Reeves, G.R., Lawrence, K.D., Lawrence, S.M., Gonzalez, J.J.: A multiple criteria approach to aggregate industrial capacity expansion. Comput. Oper. Res. **15**(4), 333–339 (1988)

Richter, L: Betriebliche Standortplanung auf regionaler Entscheidungsebene–Dissertation, Gottfried Wilhelm Leibniz Universität Hannover. In: Berichte aus dem IFA, Band 2/2017. PZH Verlag, Garbsen (2017)

Tamiz, M., Jones, D., Romero, C.: Goal programming for decision making: an overview of the current state-of-the-art. Eur. J. Oper. Res. **111**(3), 569–581 (1998)

United Nations. Department of Economic and Social Affairs, Population Division: World Urbanization Prospects–The 2014 Revision, Highlights, New York (2014)

Xu, J., Liu, Q., Lei, X.: A fuzzy multi-objective model and application for the discrete dynamic temporary facilities location planning problem. J. Civ. Eng. Manage. **22**(3), 357–372 (2016)

Zimmermann, H.-J., Gutsche, L.: Multi-Criteria Analyse–Einführung in die Theorie der Entscheidungen bei Mehrfachzielsetzungen. Springer, Berlin (1991)

Integrated Scheduling of Production and Distribution Operations with Site Selection

Eduardo Alarcon-Gerbier(✉) and Udo Buscher(✉)

Technische Universität Dresden, 01062 Dresden, Germany
{eduardo.alarcon,udo.buscher}@tu-dresden.de
https://tu-dresden.de/bu/wirtschaft/lim

Abstract. Customers' desire for ever shorter delivery times forces companies to produce and deliver customer orders in the shortest possible time. To achieve this goal, companies can follow the distributed manufacturing paradigm and try to move their production sites close to their customers, which automatically leads to a location problem that has to be solved. Another approach is to coordinate production and distribution scheduling in such a way that punctual delivery can take place at the lowest possible cost, whereby stock keeping is largely avoided. The article shows for the first time that the traditionally successively planned problems of location planning and joint production and distribution planning can be combined in an integrated approach. A multi-period MIP model is set up to coordinate both subproblems, so that the overall costs incurred can be minimized. The simultaneous planning approach is illustrated using a numerical example and then evaluated critically.

Keywords: Production and distribution scheduling · Distributed manufacturing · Optimization

1 Introduction

For some time now, the concept of Distributed Manufacturing has been attracting increasing attention. In recent decades, internationally active companies have responded to globalization by consolidating their production activities in central production facilities [18]. Undoubtedly, this enabled efficient production. As both regional and global markets have to be served in most cases, long supply chains are created, combined with long delivery times and high logistics costs. However, this often conflicts with the wishes of customers who would like their individualised products to be delivered in the shortest possible time. This is precisely where Distributed Manufacturing wants to tackle and move closer to the relevant sales markets by manufacturing in decentralised production facilities, in order to significantly shorten delivery times. However, the concept of Distributed Manufacturing is broader and is intended to represent a paradigm shift that implies a move away from mass production in centralized production facilities [13]. This concept requires organisational models for small, flexible and scalable production sites that are integrated into decentralised production networks. The aim is to

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 255–267, 2019.
https://doi.org/10.1007/978-3-030-29821-0_17

achieve extensive local production while taking local and individual customer requirements into account [17]. On-site production automatically means that the finished products no longer have to be transported over long distances, so that not only logistics costs but also CO_2 emissions are significantly reduced.

In addition to this paradigm shift, work is also being done at the operational level to meet customers' frequent requests for the fastest possible delivery. This applies in particular to the coordination between production and distribution. So far, comparatively little attention has been paid to this interface because both machine scheduling and vehicle routing planning are complex problems in themselves. The term integrated production and outbound distribution scheduling (IPODS) is used to indicate that orders are transported to external and not internal customers [10]. In the joint planning of production and distribution, two main research strands can be distinguished [19]. On the one hand, both subsystems can be directly connected with each other, so that the delivery of the products starts immediately after production. These problems are known as Integrated Scheduling of Production and Distribution Problems (ISPDP). On the other hand, it is also possible for both subsystems to be indirectly linked via a final product inventory (Integrated Scheduling of Production, Inventory and Distribution Problems (ISPIDP)). The inventory provides a greater degree of planning flexibility in both production and distribution. Despite the additional inventory costs, there are opportunities to generate more efficient solutions system-wide compared to the situation without inventory.

The aim of this contribution is to bring together the idea of decentralised or distributed manufacturing with the integrated production and outbound distribution scheduling. Formally, a location problem has to be solved together with a production and distribution scheduling problem. The paper is structured as follows. After this introduction, the Sect. 2 gives an overview of the relevant literature. In the Sect. 3, a formal description of the problem is given, followed by one MIP formulation. A numerical example is used in the Sect. 4 to illustrate the approach before the article ends with a summary in the Sect. 5.

2 Literature Review

The present contribution has its roots in the following three areas: (1) Integrated Production and Outbound Distribution Scheduling, (2) Distributed (Production) Scheduling, and (3) Location Routing Planning. The first contribution dealing with the *Integrated Production and Outbound Distribution Scheduling* (IPODS) is attributed to Potts (1980) [15] and is regarded as a pioneer for this research direction [4]. As recently as 2005, it was noted that existing approaches to integrated production, inventory and distribution planning primarily address the strategic and tactical planning level. Approaches on the operational level that deal with the scheduling of jobs are extremely rare [5]. Afterwards, the problem received a lot of attention and as a result, a large number of publications were produced. In order to systematize the literature, a first review was published in 2010, in which the problems are differentiated according to the number of customers, the type and number of vehicles and the machine configuration [4]. With regard to the main differentiation criterion 'delivery method', a distinction is made between direct delivery and batch delivery, whereby routing decisions often have to be taken into account for the latter. A second review followed in 2015, in which a primary distinction is made between whether direct delivery to the customer is made after production or whether warehousing is possible between production and distribution [19].

With regard to production, the IPODS-literature predominantly assumes that production takes place at a single location (for exceptions see for example [6,11]). The joint consideration of scheduling decisions in several geographically distributed production locations is captured with the term *Distributed (Production) Scheduling* (DS). Obviously DS proves to be more difficult than classic scheduling because scheduling decisions have to be made not only at one location but at several locations at the same time. This goes hand in hand with two dependent decisions [7]: On the one hand, customer orders must be allocated to suitable factories and on the other hand, a suitable scheduling must be carried out within the factories. This branch of research, however, concentrates exclusively on manufacturing and omits distribution decisions. Formally comparable problems are also treated in computer science under the same name. Here, however, it is a matter of executing a set of processes on different address spaces (see for example [2]).

This article shows a certain relation to *Location Routing Planning* (LRP). There are two problems to solve with LRP. On the one hand, a Facility Location Problem has to be solved on a strategic level and on the other hand a route has to be found on which the customers are supplied. The latter task, however, is on a tactical or operational level. However, the choice of location is not independent of the route decision and some studies have found, and this is not surprising, that joint planning can significantly reduce overall costs compared to successive planning [16]. While routing is characteristic of LRP, the IPODS also allows direct delivery to customers. For a detailed overview of the various facets of the LRP, the reader is referred to [14] and [9].

Although the location and route decisions are obviously interdependent, LRP is accused of mixing strategic and operational decisions. Some attempts are being made to refute these concerns by referring to lower overall costs. The approach of looking at multi-period LRPs in which the planning horizon is divided into periods is more convincing. This makes it possible to model customer demands that vary over time. To satisfy them, period-specific routes are created, while the location decision is only made once in the planning horizon. This creates a bridge from the strategic to the tactical or operative level [16]. Another very interesting approach may be seen as an extension of the approach outlined above. It suggests to map the location decision and the route decision on decoupled time scales [1]. This also opens up the possibility of closing and opening locations over time.

To the best of our knowledge, there is no contribution that simultaneously combines location planning with integrated production and outbound scheduling. In this article, we follow a multi-period approach comparable to the current developments in the field of LRP. While the location decision is only made once, the IPODS takes place in each individual period of the planning horizon. The aim of this contribution is to capture this problem mathematically in a MIP model and a numerical example is used to illustrate the interactions in the model.

3 Problem Formulation

Formally, the task can be described as follows. Let P be the set of periods indexed over $p = 1, \ldots, P$. We consider also a manufacturer that produces at M different locations to serve J decentralised customers who each place a single order with a given due date dd_{jp}. The production plants are able to produce any customer order j. Since homogeneous production sites are assumed, the processing time of an order pr_{jp} is identical across the different production sites and has to be executed without interruption. After completion (C_{jp}), each customer order is delivered directly to the customer. Due to the different locations of the production sites, a site-dependent transportation time TT_{mj} is assumed. We use D_{jp} to denote the delivery time of customer order j in period p.

Essentially, the following questions have to be answered.

1. Which sites from the pool of available sites are used for the production of customer orders (selection problem)?
2. Which customer order is manufactured at which site (assignment problem)?
3. In which order are the customer orders manufactured at the respective site (sequencing problem)?

Obviously, the answer to one question automatically influences the answer of the other. For example, the larger the number of locations, the more likely it is that an order will be produced at a location that is close to the customer. This leads to higher fixed location costs, but at the same time the transport costs and the transport times can be reduced and this in turn tends to lead to less delay costs. A larger number of locations also means that fewer orders are produced at a single location. As a result, orders will be completed and delivered earlier, which in turn will reduce the occurrence of delays. If the orders have been assigned to a site, the production sequence depends essentially on the respective due dates of the sales orders.

In the following, a cost minimization approach is chosen to satisfy customer demand with the least possible delay. The cost function consists of four cost components. First, fixed location costs FC are to be taken into account when a location is selected for the production of customer orders. Furthermore, transportation costs are incurred for the delivery of orders, which depend linearly on the transportation time (the transportation cost rate TC is measured in € per time unit). Delay costs are included as the third cost component. They are the product of the penalty cost rate PC (measured in € per time unit delay) and the customer-related tardiness T_j (measured in time units). Finally, the last cost component represents the opening costs incurred by choosing the most suitable facilities to fulfil customer orders.

We now introduce the multi-period integrated production and distribution scheduling problem with site selection. Here, the idea is to solve simultaneously the problem of selecting the optimal set of facilities and planning the machine scheduling considering multiple periods of time. The following notation is used:

Parameters:

dd_{jp} Due date of order j
FC Fixed cost per used plant per period
NL Maximum number of locations where production can take place
OC Fixed costs for opening a plant
PC Penalty costs rate for each time unit of delay
pr_{jp} Processing time of order j
Q Very large number
TC Transport costs rate per time unit
TT_{mj} Transportation time from the location of site m to the location of customer j

Variables:

C_{jp} Completion time of order j in period p
D_{jp} Delivery time of order j in period p
T_{jp} Tardiness of order j in period p
X_m Binary variable which takes the value 1 if facility m is opened at the beginning of the planning horizon
Y_{mjp} Binary variable which takes the value 1 if order j is the first order processed at site m in period p
Z_{ijmp} Binary variable which takes the value 1 if order i is processed directly before order j at site m in period p. Order $J+1$ is an artificial last order

In the following, the problem is formalized by a mixed integer program.

$$\begin{aligned}\text{Min} \sum_{m=1}^{M}\sum_{p=1}^{P}(X_m * FC) + \sum_{i=1}^{J}\sum_{j=1}^{J+1}\sum_{m=1}^{M}\sum_{p=1}^{P}(Z_{ijmp} * TT_{mi} * TC) \\ + \sum_{j=1}^{J}\sum_{p=1}^{P}(T_{jp} * PC) + \sum_{m=1}^{M}(X_m * OC)\end{aligned} \tag{1}$$

subject to

$$X_m \geq \sum_{j=1}^{J} Y_{mjp} \qquad \forall m \in M; p \in P \tag{2}$$

$$Y_{mjp} \leq \sum_{i=1}^{J+1} Z_{jimp} \qquad \forall j \in J; m \in M; p \in P \tag{3}$$

$$Z_{jjmp} = 0 \qquad \forall j \in J; m \in M; p \in P \tag{4}$$

$$\sum_{i=1}^{J} Z_{ijmp} \leq \sum_{i=1}^{J+1} Z_{jimp} \qquad \forall j \in J; m \in M; p \in P \tag{5}$$

$$\sum_{j=1}^{J+1}\sum_{m=1}^{M} Z_{ijmp} = 1 \qquad \forall i \in J; p \in P \tag{6}$$

$$\sum_{m=1}^{M} Y_{mip} + \sum_{j=1}^{J}\sum_{m=1}^{M} Z_{jimp} = 1 \qquad \forall i \in J; p \in P \tag{7}$$

$$\sum_{m=1}^{M} X_m \leq NL \qquad \forall p \in P \tag{8}$$

$$C_{jp} \geq Y_{mjp} * pr_{jp} \qquad \forall j \in J; m \in M; p \in P \tag{9}$$

$$C_{jp} \geq C_{ip} + pr_{jp} - Q * (1 - Z_{ijmp}) \qquad \forall i, j \in J; m \in M; p \in P \tag{10}$$

$$D_{jp} \geq C_{jp} + \sum_{i=1}^{J+1} (Z_{jimp}) * TT_{mj} \qquad \forall j \in J; m \in M; p \in P \tag{11}$$

$$T_{jp} \geq D_{jp} - dd_{jp} \qquad \forall j \in J; p \in P \tag{12}$$

$$T_{jp} \geq 0 \qquad \forall j \in J; p \in P \tag{13}$$

$$X_m \in \{0,1\} \qquad \forall m \in M \tag{14}$$

$$Y_{mjp} \in \{0,1\} \qquad \forall j \in J; m \in M; p \in P \tag{15}$$

$$Z_{ijmp} \in \{0,1\} \qquad \forall i \in J; j = 1,\ldots,J+1; m \in M; p \in P \tag{16}$$

In the above formulation, the objective function (1) aims at minimizing the total costs composed by facility opening costs, operational costs, transportation costs, and tardiness costs originated by delivery delay. It should be noted that production costs are not taken into account here because we are assuming homogeneous factories that need the same time to produce orders and can also produce them at the same cost. This means that the production costs are not relevant for decision-making and can therefore be omitted.

Constraints (2) guarantee that if an order j is assigned to be the first to be produced on a plant m in a period of time p, this facility has to be in operation. Constraints (3), (4) and (5) ensure in combination that for each machine m the customer order j is scheduled before the customer order i. Constraints (6) guarantee that each order i either precedes another one or is the last to be processed on a machine m. Constraints (7) force each order to be either the first processed on a machine or to follow another one. The constraints (8) limit the number of plants that can be operating per period. By (9) the completion time of the first order must be equal or greater than the corresponding processing time. Inequalities (10) determine the completion time of an order that is not the first in the sequence on a machine. This time is equal to or greater than the processing time of the job plus the completion time of its predecessor. Constraints (11) calculate the delivery time of a job by adding the corresponding travel time. Constraints (12) and (13) define the tardiness of an order as the maximum of 0 and the difference between its delivery time and its due date. Finally, constraints (14), (15), and (16) define the domain for the variables. This is a multi-period mixed integer problem that can in principle be solved with a commercial solver.

4 Numerical Example

The model introduced in the previous chapter will be illustrated using a numerical example. A problem is considered that includes six possible locations for plants and 15 customers located in different geographical places. Both the plants and the customers were randomly distributed over an area of 50 × 50 distance units (see Fig. 1).

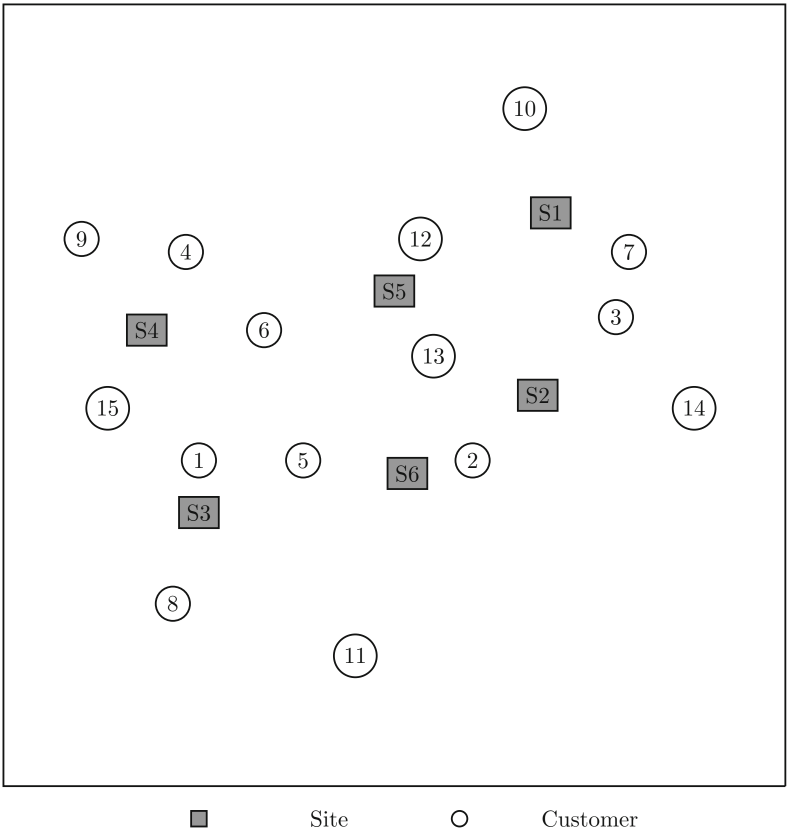


Fig. 1. Position of customers and facilities

The costs to open a site amount to €3,000 ($OC =$ €3,000). In addition, $FC =$ €4,000 are incurred per period in which a plant is used for production. The maximum number of sites that can be used per period is limited to two in case A and to three in case B. Moreover, as time-dependent transport costs are incurred, short delivery times from the plant to the customer are generally preferred. In our example, a cost rate of $TC =$ €30 is assumed for each time unit required for the delivery. The transport times between locations and customers can be found in Table 1.

Table 1. Travel time between plants and location of clients

Client	S1	S2	S3	S4	S5	S6
1	67	53	8	22	40	33
2	40	15	43	54	29	11
3	19	17	71	73	35	40
4	57	59	41	14	33	49
5	54	38	18	32	30	17
6	48	44	30	18	21	32
7	14	27	78	75	37	49
8	84	65	15	43	59	42
9	73	74	46	18	49	62
10	17	45	80	68	35	59
11	75	49	33	60	57	30
12	21	30	55	45	9	37
13	29	18	44	45	12	19
14	38	25	78	85	50	46
15	75	67	22	14	48	48

In addition to transport costs, tardiness costs are also taken into account. For each time unit of delay, a cost rate of $PC = €\,40$ is applied. In order to calculate a tardiness, the processing times of the customer orders in the individual plants as well as the agreed due dates are required in addition to the transport times. Both the processing times and the due dates are generated as random integers from a uniform distribution. Specifically, U(20,150) was assumed for the processing times and U(150, 500) for the due dates. In this example, we assume a planning time frame of three periods. Due to the longer-term nature of the location decision, the locations are determined only once at the beginning of the planning horizon. Table 2 provides the data for three periods (from P1 to P3). All other information remains the same over the periods.

IBM ILOG CPLEX version 12.7, running on a 2.6 GHz Intel Xeon CPU E5-2630 v2, was used to solve the three-periodic numerical example. The locations S2 and S3 will be opened in the first period and will remain there as assumed. The first period of case A is intended to explain the solution in more detail (cf. here and in the following also Fig. 3). While site S2 supplies six customers, site S3 only serves four customers (see Fig. 2).

Obviously, at plant S2 the customer orders are manufactured in the order C10, C7, C2, C13, C12 and C3, and at plant S3 in the order C5, C6, C14 and C11. Figure 3 also shows that the orders C6, C11, C2, C13 and C3 arrive late at the customer. Using C6 as an example, the delay can be calculated as follows: The processing of the order starts after 147 time units directly after order C5 has been completed and takes 89 time units. Since 30 time units are still necessary for the delivery, the customer receives his order C6 after 266 time units and thus with a delay of 14 time units, because the due date was 252 time units.

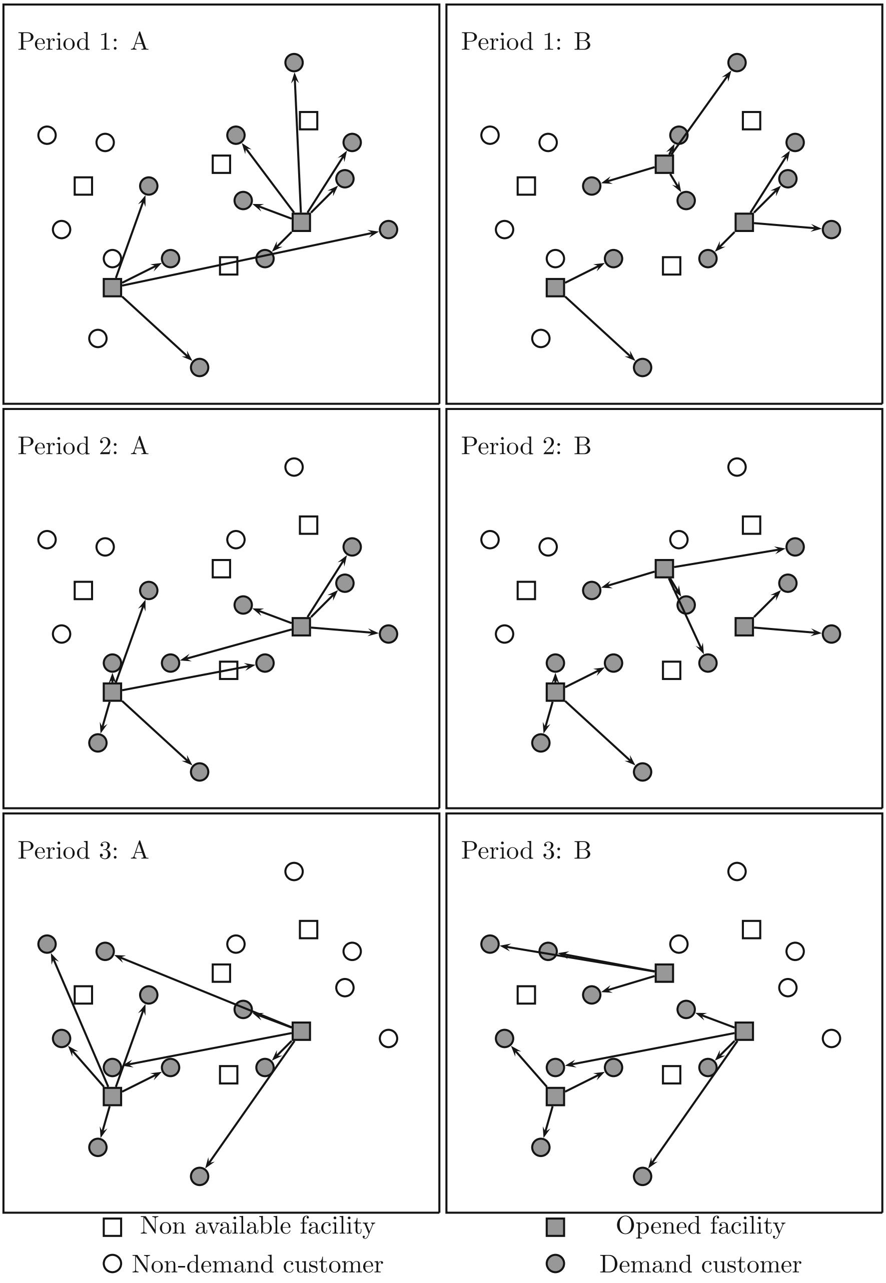


Fig. 2. Multi-Period: Position of customers and facilities

Table 2. Processing times and due dates for the numerical example

Client	P1	P2	P3
1	0	94	144
2	25	109	64
3	74	122	0
4	0	0	133
5	147	64	108
6	89	103	150
7	106	105	0
8	0	116	129
9	0	0	80
10	113	0	0
11	60	127	59
12	64	0	0
13	44	70	120
14	144	114	0
15	0	0	122

(a) Processing times

Client	P1	P2	P3
1	0	288	267
2	257	243	480
3	423	322	0
4	0	0	402
5	182	436	458
6	252	425	487
7	265	266	0
8	0	417	293
9	0	0	306
10	263	0	0
11	469	424	212
12	413	0	0
13	300	453	412
14	472	289	0
15	0	0	344

(b) Due dates

Although a different number of customer orders were assigned to the locations, it can still be seen that the total processing times of the orders are roughly the same. Since in case A the maximum number of sites to be selected was limited to two, late deliveries cannot be avoided. The total cost of the solution discussed here for case A is €79,300. They are divided as follows: €6,000 for site opening costs, €24,000 fixed costs, €26,700 transport costs and €22,600 tardiness costs. However, it should be noted that the calculation was aborted after 12 h with a gap of 76.31%. Therefore, the presented solution does not necessarily represent the optimal solution.

In case B, however, it is possible to open up to three locations. With otherwise unchanged data, this additional degree of freedom is indeed utilized and location S5 is additionally opened in comparison to case A (see again Fig. 2). Again, it can be seen that it seems advantageous to use the sites comparatively similarly in terms of processing times. In this specific case, customer orders C7, C2, C3 and C14 are now produced in S2, customer orders C5 and C11 in S3 and orders C6, C10, C13 and C12 in S5 (see Fig. 4). By the additional use of the location S5 now all orders can be delivered on time, so that no delay costs result. This means that the total costs can be reduced to €66,000. They are divided as follows: €9,000 for site opening costs, €36,000 fixed costs and €21,600 transport costs. The additional site opening costs and fixed costs are thus more than offset by the elimination of tardiness costs and reduced transport costs. As in case A, the calculation in case B was aborted after 12 h of computing time with a gap of 67.74%.

5 Summary

The aim of this contribution was to extend the integrated production and outbound distribution scheduling for the first time by the possibility of selecting from a set of possible factories those locations which are to be used for production. For this purpose,

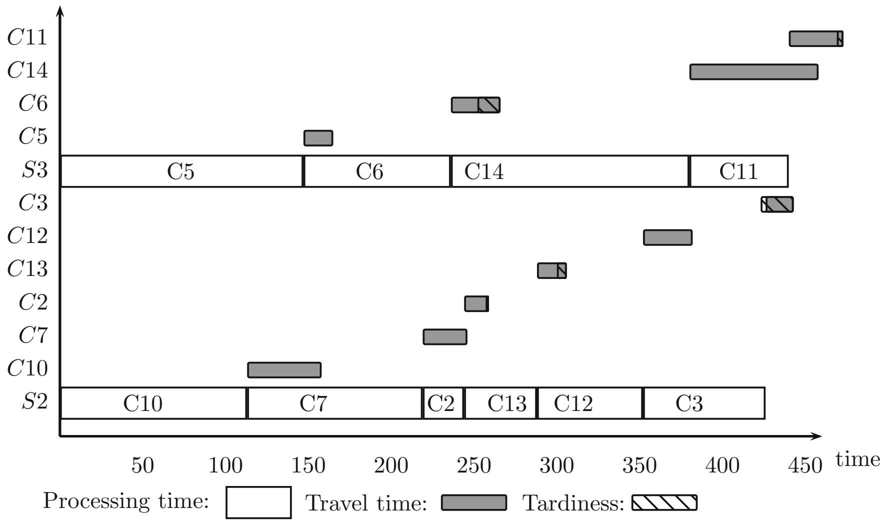


Fig. 3. Machine scheduling, case A

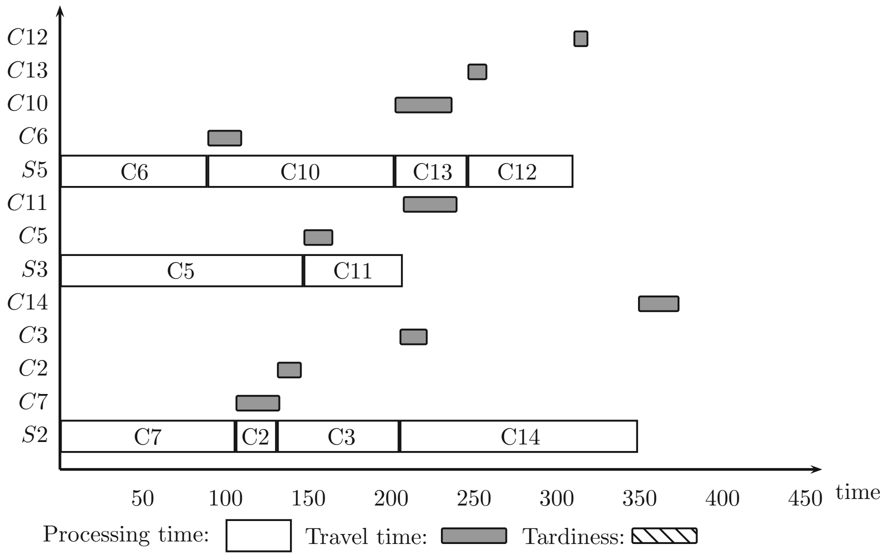


Fig. 4. Machine scheduling, case B

a MIP formulation was presented which aims at minimizing the total decision-relevant costs consisting of fixed location costs, location opening costs, transport costs and tardiness costs. As a result, the following questions will be answered: Which locations are selected and which customer orders are manufactured in which order at which location. Moreover, the numerical example essentially served to illustrate the interdependencies that occur between the choice of location on the one hand and IPODS on the other. In particular, it could be shown that the number of locations has a considerable influence on which decision-relevant costs are incurred and to what extent.

However, it must be stressed that the integrated planning approach presented is only a first conceptual proposal which can be improved in many respects. By considering several periods, the challenge of considering the more strategic choice of location and the more operational production and distribution decisions in an integrated way was at least tackled. In the future, it does not seem unrealistic that production will not necessarily take place at immobile locations, but also that it will be possible to relocate production comparatively simply. One example of this can be found in the CassaMobile project [3]. A CassaMobile is a transportable compact factory located in a twenty-foot equivalent ISO standard container. Several years earlier, the 'Factory-in-a-box' project followed a very similar approach [12]. This seems so interesting because for mobile factories the location decision no longer has to be made on a strategic but on an operational level and can therefore in fact be made together with the operational production and distribution planning.

In this context, the question also arises if the customer demand forecast for several periods is available in this level of detail in real application scenarios. The forecasting methods used certainly play a role here. Nevertheless, it is imaginable that even a comparatively rough estimate of demand could be sufficient for the selection of locations. The presented model offers various additional possibilities to extend it. The following options are examples:

– From a modelling point of view, the maximum number of sites to be used can be varied easily. In fact, the number of locations also addresses the degree of centralisation or decentralisation of the enterprise. However, the model presented does not take into account possible economies of scale that could result from a stronger bundling of customer orders at one location. This opens up the possibility of extending the model.
– First, the inclusion of site generation could be interesting. Here, the idea is to determine new locations for facilities in order to find the optimal locations. This modification requires to consider continuous space, which transforms the problem into one that is even harder to solve. Another obvious extension, instead of assuming a direct delivery, is to deliver several customer orders together with one or more vehicles.
– The simultaneous planning of production, distribution and facility location is a complex problem that can be solved with help of exact solution method only for small-sized problems. In order to solve larger instances of the problem under consideration, the development of heuristics is required.

Acknowledgement. The authors thank the Friedrich and Elisabeth Boysen-Stiftung and TU Dresden for the financial support given for this work during the third Boysen-TU Dresden-GK.

References

1. Albareda-Sambola, M., Fernández, E., Nickel, S.: Multiperiod location-routing with decoupled time scales. Eur. J. Oper. Res. **217**(2), 248–258 (2012)
2. Barroso, A.M., Leite, J.C.B., Loques, O.G.: Treating uncertainty in distributed scheduling. J. Syst. Softw. **63**, 129–136 (2002)
3. CassaMobile - A factory on the go. http://www.cassamobile.eu/
4. Chen, Z.-L.: Integrated production and outbound distribution scheduling: review and extensions. Oper. Res. **58**(1), 130–148 (2010)
5. Chen, Z.L., Vairaktarakis, G.L.: Integrated scheduling of production and distribution operations. Manag. Sci. **51**(4), 614–628 (2005)
6. Chen, Z.-L., Pundoor, G.: Order assignment and scheduling in a supply chain. Oper. Res. **54**, 555–572 (2006)
7. Chung, S.H., Chan, F.T.S., Chan, H.K.: A modified genetic algorithm approach for scheduling of perfect maintenance in distributed production scheduling. Eng. Appl. Artif. Intell. **22**, 1005–1014 (2009)
8. Drexl, A., Klose, A.: Facility location models for distribution system design. Eur. J. Oper. Res. **162**(1), 4–29 (2005)
9. Drexl, M., Schneider, M.: A survey of variants and extensions of the location-routing problem. Eur. J. Oper. Res. **241**, 283–308 (2015)
10. Fu, L.: Coordination of production and distribution scheduling. Dissertation of the University of Paris Dauphine - Paris IX (2014)
11. Garcia, J.M., Lozano, S., Canca, D.: Coordinated scheduling of production and delivery from multiple plants. Robot. Comput.-Integr. Manuf. **20**(3), 191–198 (2004)
12. Jackson, M., Wiktorsson, M., Bellgran, M.: Factory-in-a-box - demonstrating the next generation manufacturing provider. In: Mitsuishi, M., Ueda, K., Kimura, F. (eds.) Manufacturing Systems and Technologies for the New Frontier, pp. 341–346. Springer, London (2008)
13. Matt, D., Rauch, E., Dallasega, P.: Trends towards Distributed Manufacturing Systems and modern forms for their design. Procedia CIRP **33**, 185–190 (2015)
14. Nagy, G., Salhi, S.: Location-routing: issues, models and methods. Eur. J. Oper. Res. **177**, 649–672 (2007)
15. Potts, C.N.: Analysis of a heuristic for one machine sequencing with release dates and delivery times. Oper. Res. **28**, 1436–1441 (1980)
16. Prodhon, C., Prins, C.: A survey of recent research on location-routing problems. Eur. J. Oper. Res. **238**, 1–17 (2014)
17. Rauch, E., Dallasega, P., Matt, D.: Distributed manufacturing network models of smart and agile mini-factories. Int. J. Agile Syst. Manag. **10**(3/4), 185–205 (2017)
18. Srai, J., Kumar, M., Graham, G., Phillips, W., Tooze, J., Ford, S., Beecher, P., Raj, B., Gregory, M., Tiwari, M., Ravi, B., Neely, A., Shankar, R., Charnley, F., Tiwari, A.: Distributed Manufacturing: scope, challenges and opportunities. Int. J. Prod. Res. **54**(23), 6917–6935 (2016)
19. Wang, D.-Y., Grunder, O., El Moudni, A.: Integrated scheduling of production and distribution operations: a review. Int. J. Ind. Syst. Eng. **19**(1), 94–122 (2015)

Booking Limit Based Revenue Management Approaches for Customer-Value Oriented Make-to-Order Production

Nina Lohnert(✉) and Kathrin Fischer

Institute for Operations Research and Information Systems, Hamburg University of Technology, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany
nina.lohnert@tuhh.de

Abstract. In make-to-order (MTO) production, decisions are to be made about the jobs which are to be accepted and the sequence in which they are to be carried out. While in practice often rather simple rules like first-come-first-served (FCFS) are used, also strategies from the field of revenue management can be applied to achieve better results. In MTO not only the maximization of short-term profit should be focused on, but also the long-term perspective of performing good service in particular to valuable and returning customers is important. Therefore, in this work a booking-limit approach is combined with an order acceptance and scheduling model for a single machine environment to derive new strategies which take this aspect into account by defining different service levels to be strived at for the different customer segments. These strategies are tested on data settings with three customer segments. It turns out that a newly developed reversed nested booking limit approach (RNBL) leads to the best results regarding the conflicting aims of short-term profit maximization and customer satisfaction, whereas the classical partitioned booking limit (PBL) strategy is not recommendable.

Keywords: Revenue management · Make-to-order production · Booking limits · Order acceptance and scheduling · Nesting strategies

1 Introduction

These days, more and more companies use the make-to-order (MTO) principle in their production, in order to stay competitive. In MTO production, manufacturing only starts when an order has been received. This has the advantage of massively reducing the inventory of finished goods, but it leads to the difficulty that a decision has to be made regarding the orders which are to be accepted and regarding the time at which to carry them out, i.e. their scheduling. Acceptance of an order means to occupy production capacity, and the same capacity cannot be used for another, possibly more valuable order which might arrive later, and therefore has to be denied.

In this work, booking limits in combination with an order acceptance and scheduling model are used for making combined decisions on order acceptance and the subsequent scheduling. A tailored revenue management approach is developed, the basic idea of which has been presented in Lohnert and Fischer (2019). However, the

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 268–282, 2019.
https://doi.org/10.1007/978-3-030-29821-0_18

approach is modified, improved and extended here. Moreover, different booking limit strategies are developed and the results they lead to are presented.

First, some relevant background knowledge on manufacturing logistics and revenue management is briefly presented in Sects. 2 and 3. In Sect. 4, the problem definition is given, and in Sect. 5, the process of order acceptance and order scheduling is discussed and two quantitative models for these planning problems are presented. The combination of these two models and different strategies using the calculated booking limits leads to new solution procedures for the acceptance and scheduling problem. The results of the different strategies are compared in Sect. 6 using randomly generated case study data. The classical first-come-first-served (FCFS) strategy, which is often used in practice, is also included in the comparison as a reference strategy. Finally, in Sect. 7, conclusions and some directions for future research are given.

2 Theoretical Background

Production logistics is about organizing, planning, controlling and finally implementing and supervising the flow of materials and information in manufacturing companies (Gudehus 2009). In particular, capacity planning and production scheduling are important tasks in this field which will be considered by the approach presented in this work.

The major objectives that are pursued are often conflicting, as e.g. the goal of low inventory levels which leads to low costs, and the goal of service level maximization which is aimed at the maximization of customer satisfaction and the related maximization of revenue. By means of a constantly high inventory level, customers can be served quicker, but it comes with high inventory holding costs. Due to the development from a sellers' market towards a buyers' market, the customer is an important factor that is often focused on, which also leads to more variety in products to enhance customer satisfaction (Seeck 2010). As also Kersten et al. (2017) state, product individualization, cost pressure and complexity are the most important current trends in logistics, leading to changes in many companies and also in the field of production, where the individual customization of products has become more and more common.

MTO is particularly recommendable in cases where many different product varieties exist, as it helps to reduce inventories. However, this is also a disadvantage as there are no buffers, and hence any delays directly affect the customers. Still, the advantages outweigh the problems of MTO, and therefore, MTO production has become increasingly important over the past years (Barut and Sridharan 2005, Stevenson et al. 2005) and more and more companies apply this principle (Dan et al. 2018).

As these companies are confronted with uncertainty regarding the arrival of future orders, they often accept all arriving orders, including less profitable ones and more than they can actually fulfill within their limited production capacities. If employing any methods at all, they only use simple acceptance and scheduling principles like FCFS (Kalyan 2002). However, it will be shown in the following that methods from the field of revenue management can be used to allocate the scarce production resources in an improved way.

Revenue management (RM) is an approach for allocating limited, inflexible capacity in a revenue maximizing way (Guadix et al. 2010). Talluri and Van Ryzin (2005, p. 2) define RM as follows: “RM is concerned with (…) demand management decisions and the methodology and systems required to make them (…) with the objective of increasing revenues.”

RM was first introduced in the flight industry, after the deregulation of the US flight market in the late 1970s, and in the meantime has been applied in many other areas as well, e.g. in the hotel and the car rental industry. An overview over the different application areas is given by Chiang et al. (2007).

Four different techniques are included in RM, namely price discrimination, capacity control or allocation, overbooking and dynamic pricing (Klein and Steinhardt 2008). In this work, only the first two instruments will be used.

In price discrimination, different customer segments are built according to the customers’ willingness to pay. The product is offered to the customers in the different segments at different prices, but these different prices do not result from different production costs, i.e. the product is mainly identical. This kind of segmentation and price differentiation leads to larger sales and, hence, higher revenues, as the customers’ different willingness to pay is exploited (Rehkopf 2006).

Price discrimination is often combined with capacity allocation methods. There are two types of capacity control approaches, the quantity-based and the revenue-based approach. In this work, the quantity-based approach is used in which the restricted capacity is reserved for the different customer segments, resulting in so-called booking limits. Arriving customers are accepted as long as there is capacity available in the respective segment. The segments are ranked in a hierarchical manner, e.g. according to the revenue that can be achieved by selling a unit from this segment to a customer (Talluri and Van Ryzin 2005).

As with partitioned booking limits high-ranked demand may be rejected if the respective booking limit has been exhausted, although there is still capacity available in other (lower-ranked) booking limits, the nested booking limit approach has been suggested as an alternative. This approach allows the higher-ranked classes to use the capacity of the lower-ranked ones (Talluri and Van Ryzin 2005), as long as there is still capacity available in such a lower-ranked class.

3 Revenue Management in Make-to-Order Production

3.1 Requirements for Applicability and Special Characteristics

In order for RM methods to be applicable, certain requirements have to be fulfilled. These requirements can be categorized into those regarding the capacity and those regarding the demand. The requirements and their fulfillment in MTO are also discussed in e.g. Harris and Pinder (1995), Spengler and Rehkopf (2005) and Rehkopf and Spengler (2005).

Capacities have to be inflexible, i.e. limited and fixed, and perishable. This is the case, since production capacities usually cannot be expanded, at least not in the short-run and not without significant investments and as an unused hour of production capacity cannot be stored but is lost if it has not been used. The manufactured product could be stored, however. But a so-called external factor, in MTO production the customer and his/her specific order, has to be involved and hence, production in advance is impossible. Furthermore, there should be high fixed costs – which is the case in MTO as production capacities are usually expensive – and low marginal costs. The latter assumption is not fulfilled, since variable costs are not negligible in the case of MTO production. However, RM can still be applied when based on contribution margins instead of revenue, as e.g. Spengler and Rehkopf (2005) state, and this is also confirmed by other authors (e.g. Hintsches 2012, Klein and Steinhardt 2008).

Requirements regarding the demand are that demand should be heterogeneous, leading to different willingness to pay and to the possibility of market segmentation, and stochastic, i.e. fluctuating over the time (Talluri and Van Ryzin 2005). In MTO production, the ordered products, the times at which orders are placed and the due dates can differ and are not known in advance, hence these assumptions are fulfilled (Sucky 2009). Furthermore, the possibility of prior bookings should be given which leads to the decision to either accept the booking and hence block capacity in the future, or reject the booking, i.e. reserving capacity for potentially more attractive future bookings. As orders in MTO production often are accepted much earlier than they are actually carried out, this requirement is fulfilled (Harris and Pinder 1995). Additionally, historical data should be available, which also offers the option of forecasting demand (Talluri and Van Ryzin 2005). This is the case in MTO production due to the usage of, e.g., ERP systems. Moreover, there should be a standardized production program. Although in MTO production the products are not standardized, the required production capacity per product is.

Since the relevant requirements for the application of RM are fulfilled, MTO production is a promising field for the application of RM. However, there are important differences in comparison to other areas in which RM is usually applied.

As mentioned, marginal costs are not negligible here, as it is the case, e.g. in the flight industry, where an additional customer hardly leads to any additional costs. While in this case, revenue maximization can be used as an approximation of profit maximization, in MTO production the objective must be changed to maximization of the contribution margins, as stated above (Spengler et al. 2007). Since variable costs can in principle be allocated directly to products in MTO production, the cost allocation process is facilitated, in contrast to the service industry where such an allocation is usually not possible. However, it leads to an increased complexity of RM in MTO production compared to classical application areas since sophisticated cost accounting methods are required to capture the variable costs in sufficient detail (Rehkopf and Spengler 2005).

What is even more important is the fact that prices can be agreed individually with the customers. In order to develop strong customer relationships, those customers which are more important for the company, i.e. those who place orders on a regular basis, should be offered lower prices than one-time or new customers. Consequently, the value of the demand decreases with increasing (long-term) customer value which in turn results from the customer's contribution to the company's long-term monetary objectives (see e.g., von Martens and Hilbert (2011) for details on customer value). Valuable customers should receive a particularly high service level, as rejecting one of their orders might lead to a loss of many more orders in the future; therefore, it can be a reasonable objective to aim at maximization of the minimal service level of valuable customer segments (Sucky 2009).

In contrast to most of the other application areas, the planning horizon is infinite in MTO production as it is the case in the hotel industry. However, there is the particularity that while the customer is promised a certain delivery date when his order is accepted, production remains flexible until it actually gets started. Moreover, in contrast to other RM problems, delays and hence late deliveries are possible, but usually they are penalized (Barut and Sridharan 2005).

Overall, it can be concluded that there are some important differences of RM in MTO production compared to classical RM, and hence RM techniques and procedures need to be adapted accordingly in order to be applicable to MTO problems.

3.2 Literature Review

Harris and Pinder (1995) are the first to study RM in assemble-to-order (ATO) production. Kalyan (2002) states that RM in production is similar to RM in the airline industry, and concentrates particularly on the determination of bid prices. Spengler et al. (2007) point out that the complexity of RM in MTO production, especially in the steel industry, is higher than in other application areas. This might be a reason why there are only a few publications yet which study RM in MTO production.

The majority of the publications on RM in MTO considers the iron and steel industry (see Hintsches 2012, Hintsches et al. 2010, Rehkopf and Spengler 2005, Spengler and Rehkopf 2005, Spengler et al. 2007). In most cases, revenue-oriented capacity allocation is used, e.g. a capacity allocation based on opportunity costs or on bid prices (see e.g., Talluri and Van Ryzin 2005 for details on bid prices). Bid prices are also used for MTO problems in other industries, see, e.g., Elimam and Dodin (2001), Kalyan (2002) and Herde (2018).

A first contribution considering capacity allocation problems in MTO production is Balakrishnan et al. (1996). Barut and Sridharan (2005) develop a heuristic approach for making capacity-allocations and hence order acceptance decisions, while Kuhn and Defregger (2004) and Defregger and Kuhn (2007) use a Markov-chain based approach. Another approach using quantity-oriented capacity planning is developed by Kumar and Frederick (2007) for the home construction industry.

Sucky (2009) derives booking limits for order acceptance decisions and studies different effects of profit and service level maximization as well as concepts for deriving a compromise for these two objectives. Below, a quantity-based approach is presented which includes ideas from Sucky (2009) and builds up on the capacity allocation model by Klein and Steinhardt (2008) and on the order acceptance and scheduling model by Thevenin et al. (2016), but extends these approaches in different directions.

4 Problem Definition

In this work, the order acceptance and scheduling process of a make-to-order driven company in a single machine environment is studied. The materials which are required for production have to be ordered from suppliers. Two different types of delivery are available: The standard delivery and an express delivery which is faster but comes with extra costs. These costs also depend on the size of the order. Furthermore, for each job a due date and a deadline are given. If a job is finished after the due date, it is delayed and the customer will not pay the full price, e.g. because of contract penalties. However, no job may be completed later than its deadline which is also known.

It is assumed that the company under consideration has been operating in the market for a while already, and therefore has a partly known customer base. The customers are assigned to different segments. Customers who have a long-term framework contract with the company are assigned to segment A. They have a high customer value for the company as they are returning customers, but due to their contract they get discounts on the price they pay per order, leading to a lower (short-term) value of their demand for the company. If a customer does not have such a contract but has already ordered sometimes at this company, the customer is assigned to segment B. These customers do not get any special conditions. Hence, the value of a customer from segment B can be described as "medium", and the same is true for the respective demand value. Finally, new customers are assigned to segment C. Since their value for the company cannot be evaluated yet, the customer value is considered to be low. However, the demand value is the highest of the customer segments, since the new customers' negotiation power is weakest and therefore they have to pay the highest prices. Moreover, also the orders as such are grouped into different categories according to their lead time, i.e. into short, medium and long orders.

The planning problem at hand is to decide which orders to accept, and when and in which sequence to carry them out in order to maximize profit while violating the aspired service levels for the different customer groups as little as possible.

5 Solution Approach

With the approach presented below, the order acceptance and the ensuing job scheduling is optimized as follows: A preliminary decision will be made by means of the booking limits which result from the capacity allocation model presented in Sect. 5.1. If the order is preliminarily accepted, it will be determined by the order acceptance and scheduling model (OAS) (see Sect. 5.2) if the order can actually be fulfilled before its deadline. If the order is to be accepted according to the OAS-model, the customer is informed accordingly. Finally, in practice also the customer can withdraw his order, e.g. if he is not satisfied with the offered delivery time. However, this aspect is ignored in the following. This order acceptance process will be conducted in specified time intervals, e.g. once per day, as new orders arrive at the company over time. Figure 1 shows this process for a single job j.

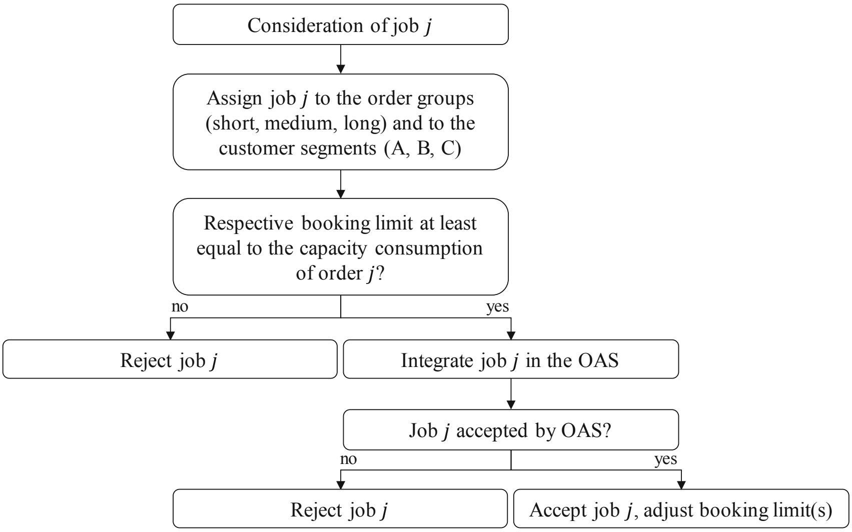


Fig. 1. Order acceptance process

5.1 Capacity Allocation Model

Using a capacity allocation model, which is based on Klein and Steinhardt (2008), optimal partitioned capacity-oriented booking limits can be determined in a first step.

Two sets are required for the model formulation. The set K contains the customer segments k, whereas Q includes the order groups q. It is important here that the combination of one element of each set is a product in the sense of RM. For each such product a booking limit x_{qk} will be determined which indicates the capacity (measured in time units) which is reserved for the respective product. For each product a demand forecast, denoted by DF_{qk}, is given for the planning horizon. Since the actual demand is unknown, it is assumed that the forecast DF_{qk} corresponds well to the actual demand and therefore reflects the orders to be expected. By the parameters aCM_{qk}, the average contribution margins for the products are given. These values are derived from historical data.

Within the planning horizon, there is a fixed production capacity CAP available on the machine. The capacity consumption of a product is specified by the order group q and mLT_q denotes the maximum lead time for this group. It is important to consider the maximum lead time (and not the average) since with this approach the real (at this point unknown) lead time cannot be underestimated and hence, the determined solution will never be infeasible.

In contrast to Klein and Steinhardt (2008), here customer segments and the corresponding aspired service levels α_k play an important role in the capacity allocation model. The service levels α_k indicate the percentage of orders of the respective customer segment k that should be accepted. Since the maximization of contribution margins does not align with the maximization of the service level (Sucky 2009), a compromise solution has to be found for these conflicting aims. The approach taken here is an objective weighting approach, where deviations dev_k from the aspired service levels α_k are weighted with penalty costs $Pdev_k$ and contribute negatively to the objective function which otherwise consists of contribution margin maximization.

The capacity allocation model for the determination of optimal partitioned booking limits x_{qk} can be formulated as follows:

$$maximize \sum_{q\in Q}\sum_{k\in K} aCM_{qk}\cdot\frac{x_{qk}}{mLT_q} - \sum_{k\in K} Pdev_k \cdot dev_k \quad (1)$$

$$\sum_{q\in Q}\sum_{k\in K} x_{qk} \le CAP \quad (2)$$

$$\frac{x_{qk}}{mLT_q} \le DF_{qk} \qquad \forall q\in Q, k\in K \quad (3)$$

$$\sum_{q\in Q}\frac{x_{qk}}{mLT_q} \ge (\alpha_k - dev_k)\cdot\sum_{q\in Q} DF_{qk} \qquad \forall k\in K \quad (4)$$

$$dev_k \ge 0 \qquad \forall k\in K \quad (5)$$

$$x_{qk} \ge 0 \qquad \forall q\in Q, k\in K \quad (6)$$

As stated above, the objective function (1) maximizes the sum of the contribution margins minus the penalty costs for deviations from the aspired service levels. Note that the quotient of the booking limit x_{qk} and the maximum lead time mLT_q gives the product-oriented booking limit (number of product units to be produced). Constraint (2) guarantees that the sum of the booking limits x_{qk} does not exceed the available capacity CAP. With restrictions (3), it is ensured that the product-oriented booking limits do not exceed the demand forecasts DF_{qk}. By means of constraints (4), deviations dev_k from the aspired service levels α_k are captured. Finally, constraints (5) and (6) guarantee the non-negativity of the variables dev_k and x_{qk}.

5.2 Order Acceptance and Scheduling Model

The OAS-model is based on the work of Thevenin et al. (2016). However, the model is adapted to the specific planning situation as, for example, the aspired service levels α_k of the different customer segments k are taken into account. Since there is a link between this model and the capacity allocation model, some of the notation from the capacity allocation model is also used in the OAS-model.

All arriving orders j are included in the set J. By means of the indicator parameters L_{kj} and $\overline{L}_{qj}$, it is captured to which customer segment k and to which order group q job j belongs. If job j originates from customer segment k, L_{kj} is set to 1, and otherwise to 0. In the same manner it can be captured whether the job j is from order group q ($\overline{L}_{qj} = 1$) or not ($\overline{L}_{qj} = 0$).

Furthermore, there are several subsets of J. It is important to remember that the OAS-model has to be solved several times (e.g. once per day in the planning horizon). Hence, decisions which have been made in the past must be recorded and remain valid. Therefore, all jobs j which have already been accepted have to be added to the subset H

of accepted orders, and all jobs j which have been rejected either by the booking limits or by the OAS-model have to be added to the subset A of rejected orders. Hence, only the jobs j which arrive at the day of the execution are not assigned to any subset of J. Moreover, the sets $J0$ and Jn contain one dummy order each, one for the beginning and one for the end of the scheduling.

The acceptance decision of a job j is modeled by binary variables z_j ($z_j = 1$ if order j is accepted). Furthermore, for the decisions regarding the acceptance and scheduling of jobs some scheduling related parameters are necessary. The estimated lead time (including a buffer) of job j is given by the parameter LT_j. Furthermore, for each job j a regular release date R_j and an early release date $\bar{R}_j$ (resulting from material availability, i.e. the type of delivery which is chosen) is given. D_j denotes the due date and $\bar{D}_j$ represents the deadline of a job j. Both dates are defined by the customer. The variables b_j and f_j contain the times of the beginning and the completion of the respective job's production. Since the OAS-model includes sequence-dependent setup times, the setup time from order j to order i is represented by parameter ST_{ji} and the setup costs are given by parameter SC_{ji}. The sequence of orders can be determined by the binary variable y_{ji} which equals 1, if order j directly precedes order i. The individual sales prices minus several directly attributable costs are denoted by CM_j. Moreover, there are penalty costs. PE_q give the penalty costs for early start of production (per time unit) while the time is captured by variables e_j, and likewise the penalty costs PT_k are given for delayed completion (per time unit), and the number of time units a job is delayed are recorded by variables t_j. Finally, by the penalty cost parameter PU any unused production capacity (variable u) is punished, since a high capacity utilization of the perishable capacity is strived for.

Thus, the following order acceptance and scheduling model results:

$$\begin{aligned} \text{maximize} \sum\nolimits_{j \in J} CM_j \cdot z_j - \sum\nolimits_{j \in J0 \cup J} \sum\nolimits_{i \in J \cup Jn: i \neq j} SC_{ji} \cdot y_{ji} \\ - \sum\nolimits_{q \in Q} \sum\nolimits_{j \in J} PE_q \cdot \bar{L}_{qj} \cdot e_j - \sum\nolimits_{k \in K} \sum\nolimits_{j \in J} PT_k \cdot L_{kj} \cdot t_j \\ - \sum\nolimits_{k \in K} Pdev_k \cdot dev_k - PU \cdot u \end{aligned} \tag{7}$$

$$f_j = b_j + LT_j \qquad \forall j \in J \tag{8}$$

$$t_j \geq f_j - D_j - \bar{D}_j \cdot \left(1 - z_j\right) \qquad \forall j \in J \tag{9}$$

$$e_j \geq R_j - b_j - \overline{D}_j \cdot \left(1 - z_j\right) \qquad \forall j \in J \tag{10}$$

$$b_j \geq f_i + ST_{ij} \cdot y_{ij} - \overline{D}_i \cdot \left(1 - y_{ij}\right) \qquad \forall j \in J \cup Jn, i \in J0 \cup J : i \neq j \tag{11}$$

$$b_j \geq \overline{R}_j \qquad \forall j \in J \tag{12}$$

$$f_j \leq \bar{D}_j \qquad \forall j \in J \tag{13}$$

$$z_j - \sum_{i \in J \cup Jn: i \neq j} y_{ji} = 0 \qquad \forall j \in J0 \cup J \tag{14}$$

$$z_j - \sum_{i \in J0 \cup J: i \neq j} y_{ij} = 0 \qquad \forall j \in J \cup Jn \tag{15}$$

$$\sum_{j \in J \setminus A} L_{kj} \cdot z_j \geq (\alpha_k - dev_k) \cdot \sum_{j \in J} L_{kj} \qquad \forall k \in K \tag{16}$$

$$b_{n+1} - \sum_{j \in J} \bar{R}_j \cdot y_{0j} - \sum_{j \in J} LT_j \cdot z_j - \sum_{j \in J0 \cup J}^{n} \sum_{i \in J \cup Jn: i \neq j} ST_{ji} \cdot y_{ji} = u \tag{17}$$

$$z_j = 1 \qquad \forall j \in H \cup J0 \cup Jn \tag{18}$$

$$z_j = 0 \qquad \forall j \in A \tag{19}$$

$$b_{n+1} \leq CAP \tag{20}$$

$$dev_k \geq 0 \qquad \forall k \in K \tag{21}$$

$$b_j, e_j, f_j, t_j \geq 0 \qquad \forall j \in J \tag{22}$$

$$u \geq 0 \tag{23}$$

$$y_{ji} \in \{0, 1\} \qquad \forall j \in J0 \cup J, i \in J \cup Jn : i \neq j \tag{24}$$

$$z_j \in \{0, 1\} \qquad \forall j \in J \tag{25}$$

In the objective function (7) the sum of the contribution margins of all accepted orders minus the setup and the different penalty costs is maximized. With constraints (8), the beginning of production and the completion time of a job j are set into a fixed relation, depending on the time required for job j. If the completion of a job j is delayed or the start of production is early, this is captured by constraints (9) and (10). Constraints (11) and (12) specify the beginning of the production. For the start of production of a job j, the preceding job i has to be completed and the setup must have been carried out. Furthermore, constraints (12) guarantee that the processing of a job j cannot be started before the necessary materials have been delivered. By means of constraints (13) it is ensured that the deadline $\bar{D}_j$ is the latest possible time of completion of job j. Constraints (14) and (15) guarantee that every accepted order has exactly one preceding and exactly one subsequent job.

Constraints (16) are the service-level-restrictions. Any deviations from the service levels α_k will be determined through these constraints. By means of restriction (17), any unused production capacity will be identified and penalized in the objective function (7). The beginning b_{n+1} of the dummy order $n+1$, i.e. the end of the production time, is compared with the early release date $\overline{R}_j$ of the first scheduled order j. By subtracting the sum of all lead times LT_j of the accepted orders and the sum of the setup times ST_{ji}, the result gives the unused capacity u.

By constraints (18) it is ensured that every order from subset H and the dummy orders will be accepted. Consequently, every order from subset A is rejected by means of constraints (19). Restriction (20) limits the planning horizon to its production capacity CAP. Finally, constraints (21) to (25) define the variables' definition areas.

5.3 Booking Limit Acceptance Strategies

The booking limits x_{qk} can be used in various ways in order to decide about the (preliminary) acceptance of bookings. One of them is the direct usage of the partitioned booking limits x_{qk} (PBL) calculated by the capacity allocation model. However, this has the disadvantage that bookings in higher valued classes cannot use the capacity of lower valued classes. As mentioned above, the nesting of booking limits is a possibility to get improved results. Three different nesting strategies will be explained in the following. For all nesting options considered it can be assumed that a short order which requires less time is more profitable per time unit than a medium one, and this in turn is more profitable than a long order.

One option is the nesting based on the demand value (NDV). For this approach, the products have to be ranked in a hierarchical manner according to their average contribution margins aCM^{*}_{qk} per time (and hence capacity) unit. Here, the superordinate criterion for the demand value of the product is the customer segment k, and the subordinate criterion is the order group q. Hence, a short order from a C-customer has the highest demand value, whereas a long order of from segment A has the lowest demand value. Consequently, a short order from a C-customer can use the booking limits of all other classes which gives preference to the – in terms of customer value less valuable – C-customers.

Another option is to nest the booking limits only within the customer segments (NWCS) in order to restrain, e.g., C-orders from using class A capacity. Thus, e.g. a short order of customer segment C can only access the booking limits $x_{medium,C}$ and $x_{long,C}$ and not those of the other customer segments.

Furthermore, so-called reverse nested booking limits (RNBL) can be used in order to pursue a higher customer satisfaction of the crucial customers, i.e. those from segment A. Therefore, again the customer segment k is the superordinate criterion, but with A as the highest ranked and C as the lowest ranked customer segment. The subordinate criterion is again the order group q. So in this case, a short order from an A-customer can use the booking limits of all other classes.

The nesting and the adjustments of the booking limits follow the rules of the standard nesting (Talluri and Van Ryzin 2005) and hence the nested booking limits can be determined by a summation of the partitioned booking limit x_{qk} itself and the booking limits of all accessible lower-ranked products. For a detailed explanation of the standard nesting process and the necessary adjustments in the case of the NDV strategy, see Lohnert and Fischer (2019).

If an order is accepted, the booking limit/s has/have to be adjusted by the actual capacity usage. Hence, the real (at this point known) lead time LT_j plus 5% of the maximum lead time mLT_q, which is used as an approximation of the sequence-dependent setup times, is subtracted from the respective booking limit/s. Further orders can only be accepted as long as the respective booking limit is sufficiently large.

6 Numerical Studies

Various case studies have been conducted in order to compare the four different booking limit acceptance strategies. Furthermore, FCFS, the most common acceptance approach in practice (Kalyan 2002), is considered for comparison. For all case studies, order data were randomly generated, but it was always ensured that the demand exceeds the supply and that the share of the orders of customer segment A makes up about 50% of the orders, whereas the share of orders originating from customer segment B contributes around 30% and the remaining orders are from new customers. Due to space limitations, the full data sets of these case studies and the complete set of rules used for their generation cannot be given here, but the data are available from the authors on request.

It is assumed that the demand forecast, hence the amount of expected orders of the different products, corresponds to the actual demand, i.e. the forecast is assumed to be perfect. However, this does not mean that the actual amount of capacity required is the same as predicted since the actual lead time LT_j might be smaller than the maximum lead time mLT_q. All penalty cost coefficients are set to fixed values, such that penalty costs which relate to a customer segment k are set to higher values, since they indicate that the customer is directly affected, compare to those which depend only on the order group q. $Pdev_k$ have even higher values than PT_k. The value of PU results from the objective function value of the capacity allocation model divided by the available capacity CAP, which is one month, i.e. 20 days, in all case studies.

For the different acceptance strategies, a recurring pattern over the various case studies was clearly visible. In the following, the results of one exemplary selected case study will be presented in order to illustrate the major findings. This case study includes 33 orders which are randomly generated according to the above-mentioned rules and arrive randomly during the planning horizon. Figure 2 shows the results of the different order acceptance strategies.

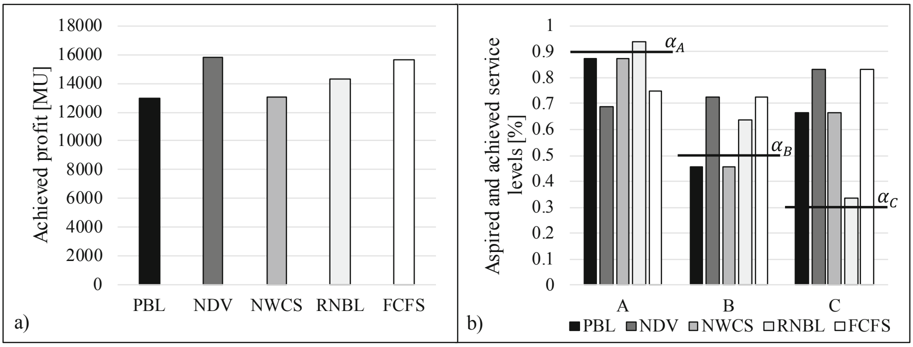


Fig. 2. Results of acceptance strategies in comparison: (a) profit, (b) service levels

As can be seen in Fig. 2a, the highest profit, i.e. the sum of the contribution margins, calculated only including actual variable costs and excluding penalty costs

$Pdev_k$ and PU, can be achieved by using the NDV strategy. With roughly 1% less, the second highest value in terms of profit can be achieved by using FCFS. This is quite a good performance of FCFS, taking into account that the resulting scheduling of the orders is not optimal. In the other case studies, FCFS also leads to the second highest profit, but with a larger difference compared to the highest value which is always achieved with NDV. With the RNBL strategy, a somewhat lower value can be gained. PBL and NWCS both attain the same result and in all cases the lowest profit.

However, it is crucial to also compare the achieved service levels for the different customer segments (Fig. 2b), since these different segments are of different relevance for the future long-term success of the company. As pointed out above, it is important to satisfy A-customers in order to fulfill the framework contract and therefore stabilize the customer base. New customers (segment C) are less important compared to the other segments. The aspired service levels α_k are therefore set to 90% for A-customers, 50% for B-customers and 30% for C-customers.

The best result regarding the service levels can be achieved with the RNBL strategy, as the service level for customer segment A is above 90% and also the aspired service levels for segment B and C are achieved by the RNBL strategy. The second highest service level for segment A can be achieved by PBL and NWCS. However, customer segment B has a rather low service level when those acceptance strategies are used, in contrast to RNBL which leads to quite a good value also for customer segment B, whereas it does not accept many customers from segment C. While in this case study, PBL and NWCS lead to the exactly same result, in other cases NWCS performed better in terms of profit achieved, without decreasing the service level compared to PBL. Hence, NWCS dominates PBL.

NDV and FCFS achieve the highest profit but they reach those good results because they accept most of the orders of customer segment C (high demand value, but low customer value). Especially FCFS does not take any service level requirements into account at all. Hence, FCFS is not a useful strategy when this long-term objective is of importance, but it performs rather well when only the profit is considered, as it leads to a high usage of available capacities.

It should be noticed that capacity utilization differs for the different strategies. The sum of all partitioned booking limits x_{qk} always equals the available capacity CAP but since the actual lead time LT_j might be smaller than the maximum respective lead time mLT_q, it is possible that the booking limits are not completely exhausted during the planning horizon. Thus, capacity can remain unused. In the case study considered here, the highest level of capacity utilization (98%) is reached by FCFS, closely followed by the acceptance approach NDV. With RNBL, a capacity utilization of 91% was reached. PBL and NWCS lead to a capacity usage of 80%.

Overall, the different case studies showed that the trade-off between profit maximization and the maximization of the service levels of the important customers can be optimized when the RNBL acceptance strategy is used. This strategy achieves the highest service level for customers in segment A and at the same time enables a rather high profit, while making good use of the available capacity.

7 Conclusions and Outlook

Since the customization of products leads to an enormous increase of variant diversity, more and more companies use the concept of MTO production in order to stay competitive. However, due to stochastic and volatile demand, optimal order acceptance decisions are rather difficult to make. Therefore, the application of appropriate RM strategies may help in facilitating and optimizing these decisions.

In this work, a capacity allocation model and an OAS-model are combined to develop different strategies for making order acceptance and scheduling decisions. The application of those strategies and the classical FCFS approach to different case studies shows that a new strategy using RNBL leads to the best compromise between the maximization of total short-term profit and the maximization of the service level of crucial customers which will contribute to the long-term success of the company. Hence, the RNBL approach is particularly suitable for MTO problems.

In future work, a larger number of cases has to be studied within a simulation study framework to confirm that the results stated above are independent of the underlying data. Moreover, the relation between different cost parameters, especially the penalty costs for the deviation from the aspired service levels, has to be studied in sensitivity analyses in order to investigate their influence on the acceptance of jobs. Furthermore, the performance of the different booking limit acceptance strategies should be tested when demand differs significantly from its forecast. Moreover, it should be studied how the long-term effects of the achieved service levels, in particular with respect to customers with low demand but high customer value, could be measured. In the approach presented here it is assumed that a higher service level will lead to higher long-term profits; however, this might be studied empirically and an appropriate approach for modeling this aspect might be built.

Also in the scheduling process more aspects could be considered like for instance more machines and the possibility of using overtime. Hence, RM in MTO is a very promising field for future research.

References

Balakrishnan, N., Sridharan, V., Patterson, J.W.: Rationing capacity between two product classes. Decis. Sci. **27**(2), 185–214 (1996)

Barut, M., Sridharan, V.: Revenue management in order-driven production systems. Decis. Sci. **36**(2), 287–316 (2005)

Chiang, W.-C., Chen, J.C., Xu, X.: An overview of research on revenue management: current issues and future research. Int. J. Revenue Manage. **1**(1), 97–128 (2007)

Dan, B., Gao, H., Zhang, Y., Liu, R., Ma, S.: Integrated order acceptance and scheduling decision making in product service supply chain with hard time windows constraints. J. Industr. Manage. Optim. **14**(1), 165–182 (2018)

Defregger, F., Kuhn, H.: Revenue management for a make-to-order company with limited inventory capacity. OR Spectrum **29**(1), 137–156 (2007)

Elimam, A.A., Dodin, B.M.: Incentives and yield management in improving productivity of manufacturing facilities. IIE Trans. **33**(6), 449–462 (2001)

Guadix, J., Cortés, P., Onieva, L., Muñuzuri, J.: Technology revenue management system for customer groups in hotels. J. Bus. Res. **63**(5), 519–527 (2010)

Gudehus, T.: Comprehensive Logistics, 1st edn. Springer, Heidelberg (2009)

Harris, F.H.deB., Pinder, J.P.: A revenue management approach to demand management and order booking in assemble-to-order manufacturing. J. Oper. Manag. **13**(4), 299–309 (1995)

Herde, F.: Revenue-Management-Ansatz für eine Annahmesteuerung kundenspezifischer Regenerationsaufträge komplexer Investitionsgüter. Springer, Wiesbaden (2018)

Hintsches, A.: Dynamische Kapazitätssteuerung bei kundenindividueller Auftragsproduktion in der stahlverarbeitenden Industrie, 1st edn. Gabler Verlag, Wiesbaden (2012)

Hintsches, A., Spengler, T.S., Volling, T., Wittek, K., Priegnitz, G.: Revenue management in make-to-order manufacturing: case study of capacity control at ThyssenKrupp VDM. Bus. Res. **3**(2), 173–190 (2010)

Kalyan, V.: Dynamic customer value management: asset values under demand uncertainty using airline yield management and related techniques. Inf. Syst. Front. **4**(1), 101–119 (2002)

Kersten, W., Seiter, M., von See, B., Hackius, N., Maurer, T.: Chancen der digitalen Transformation: Trends und Strategien in Logistik und Supply Chain Management. DVV Media Group GmbH (2017)

Klein, R., Steinhardt, C.: Revenue Management. Springer Verlag (2008)

Kuhn, H., Defregger, F.: Revenue Management in der Sachgüterproduktion. WiSt-Wirtschaftswissenschaftliches Studium **33**(5), 319–324 (2004)

Kumar, S., Frederick, J.L.: Revenue management for a home construction products manufacturer. J. Revenue Pricing Manag. **5**(4), 256–270 (2007)

Lohnert, N., Fischer, K.: Einsatz des Revenue Managements in der Make-to-Order-Produktion. In: Schröder, M., Wegner, K. (eds.) Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains: Festschrift für Wolfgang Kersten zum 60. Geburtstag, pp. 235–262. Gabler Verlag (2019)

Rehkopf, S.: Revenue Management-Konzepte zur Auftragsannahme bei kundenindividueller Produktion: Am Beispiel der Eisen und Stahl erzeugenden Industrie. Springer Verlag, Wiesbaden (2006)

Rehkopf, S., Spengler, T.: Revenue management in a make-to-order environment. In: Operations Research Proceedings 2004, pp. 470–478. Springer, Heidelberg (2005)

Seeck, S.: Erfolgsfaktor Logistik. Gabler Verlag (2010)

Spengler, T., Rehkopf, S.: Revenue Management Konzepte zur Entscheidungsunterstützung bei der Annahme von Kundenaufträgen. Zeitschrift für Planung & Unternehmenssteuerung **16**(2), 123–146 (2005)

Spengler, T., Rehkopf, S., Volling, T.: Revenue management in make-to-order manufacturing – an application to the iron and steel industry. OR Spectr. **29**(1), 157–171 (2007)

Stevenson, M., Hendry, L.C., Kingsman, B.G.: A review of production planning and control: the applicability of key concepts to the make-to-order industry. Int. J. Prod. Res. **43**(5), 869–898 (2005)

Sucky, E.: Revenue Management bei Auftragsfertigung – Ein Ansatz zur Kapazitätssteuerung. In: Voß, S., Pahl, J., Schwarze, S. (eds.) Logistik Management, pp. 77–96. Physica Verlag (2009)

Talluri, K.T., Van Ryzin, G.J.: The Theory and Practice of Revenue Management. Springer, Heidelberg (2005)

Thevenin, S., Zufferey, N., Widmer, M.: Order acceptance and scheduling with earliness and tardiness penalties. J. Heuristics **22**(6), 849–890 (2016)

von Martens, T., Hilbert, A.: Customer-value-based revenue management. J. Revenue Pricing Manag. **10**(1), 87–98 (2011)

An IP Formulation for a Machine-Sequencing Problem to Minimize Job Deviations and Set-Ups

Tommy Schultz(✉)

BMW Group Plant Leipzig, BMW Allee 1, 04349 Leipzig, Germany
tommy.schultz@bmw.de

Abstract. The maximum deviation machine-sequencing problem with set-ups is a short-term decision problem in which a sequence of plastic outer-shell car parts are to be determined for launching in a paint shop. The decision is made under two objectives. On the one hand, the deviation between the original job number in the called off sequence by the customer and the position on which the job is painted should be minimized. On the other hand, the number of set-ups that occur when changing the colour in the paint shop should also be minimized. The situation is taken from a real-world situation of a supplier in the German automotive industry. Due to the large number of authentic influences, the problem is described in detail and classified according to existing characterizations of scheduling problems. An IP formulation is presented that is solved by the ϵ-constraint method. Finally, an example is explained and computational results discussed.

Keywords: Machine-sequencing problem · Paint shop · ϵ-constraint method

1 Introduction

Methods for sequencing and scheduling should ensure the optimal utilization of scarce resources through production planning. Due to their high practical relevance, the research dates back to the 1950s [1]. If there is a machine in a production process that represents a bottleneck for the overall process, efficiency can be increased by applying the single-machine sequencing problem. A sequence of orders for production on the machine must be identified so that one or more objectives, such as the total machining time, can be optimized. The reduction of set-up costs is also a possibility to increase efficiency, as Wilbrecht and Prescott [2] showed.

If there is a sequence of jobs $1 \ldots J$ given that the supplier has to deliver, the task is to rearrange the orders so that he can optimize production. This problem is called the *resequencing problem*. The terms *order* and *job* are used synonymously in this article and refer consistently to a component to be delivered

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 283–297, 2019.
https://doi.org/10.1007/978-3-030-29821-0_19

in a sequence called off by the OEM. Resequencing problems have often been studied in cases of resequencing orders for final assembly or paint shops [3].

This article discusses the problem of sequencing during order release in a paint shop of a first-tier supplier for plastic outer-shell parts in the German automotive industry. These suppliers must deliver the finished work pieces just-in-sequence (JIS) and just-in-time (JIT) to an original equipment manufacturer (OEM). The order sequence is transmitted from the customer to the supplier four to six days earlier [4]. Along with various objectives, this sequence must be changed in order to increase manufacturing efficiency and the supplier's competitiveness.

First, the problem from practice and its objectives will be described, and an overview of the research situation will be given. Based on this, an IP formulation that meets the real-world requirements will be presented and described in detail. In addition, an example will be shown to explain the functionality, and the computational results will be discussed. Finally, needs for further research are noted.

2 Problem Description from Practice

The following section describes the manufacturing process and the focused decision problem. The description will also classify the properties of the existing problem according to the classification $\alpha|\beta|\gamma$ by Graham et al. [5], where α represents the machine environment, β the job characteristics and γ the objective function.

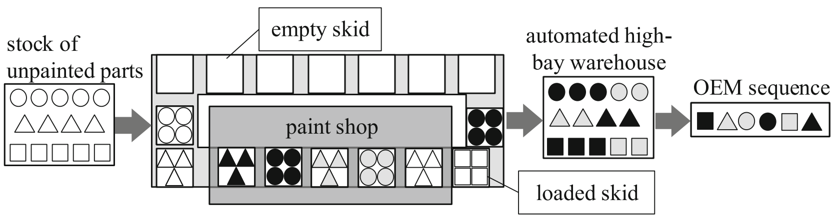


Fig. 1. Paint-shop process

The manufacturing process of a first-tier supplier for plastic outer-shell parts, such as bumpers and sill panels, consists essentially of the production of the raw parts by injection moulding and the subsequent painting of these parts, as Fig. 1 shows. Three layers of lacquer are applied to these parts in the paint shop: primer, base coat and clear coat. For each layer, a separate process section exists where the application of paint is done automatically by industrial robots. The process section for the base coat is the bottleneck of the whole paint shop. Therefore, the paint shop can be regarded as one machine: $\alpha_1 = 1$. Transportation is conducted by a chain conveyer with a constant speed so that processing times for jobs are equal. Therefore, it is impossible to split or interrupt jobs. An interruption in the paint application process would also lead to defective parts. In addition, for efficiency, no idle times are allowed on the machine.

The raw materials for painting are provided from a warehouse, and all of them are continuously available. For the paint process, they have to be placed on special holders mounted on a skid. A single skid must always be equipped with a defined number of the same parts, all painted the same colour. Currently, more then 5 different types of skids are available, and the application of 15 different colours is possible. Regarding the notation by Graham et al. [5], the jobs can be characterized as $p_j = 1$ because the process times are the same for each job due to the constant speed of the conveyer. In addition, no idle times for the paint shop have to be taken into account. For these reasons, it is sufficient to plan a sequence of jobs, and there is no need to create a schedule.

The components are inspected and stored in a warehouse after the painting process. The sequence called off by the OEM is supplied from the available parts in this warehouse. The time available is less than seven hours between the retrieval of the parts from storage and mounting them at the OEM's assembly line. The warehouse, therefore, has the function of compensating for malfunctions in the paint shop in order to guarantee the supply of the OEM. Figure 2 shows the stock in the warehouse and the sequence called off by the OEM. The first six orders and the eighth order of the sequence can be delivered immediately from the warehouse, so orders labelled from one to eight must be produced next in paint-shop.

The focused decision problem is to determine the sequence in which these jobs are released to the paint shop. The orders must be assigned to their corresponding skid types under the condition that only orders of the same colour are on a skid. The sequence of the skids is freely selectable, and there are no restrictions regarding the availability of skids. As painting sequence option one of Fig. 2 shows, jobs one and three are assigned to skid one, jobs four and seven to skid two, and so on.

The assignment decision is made under two objectives. On the one hand, the sequence in which the components are painted must correspond as closely as possible to the sequence called off by the OEM; an explanation follows later. On the other hand, the number of colour set-ups should be reduced. A colour set-up occurs whenever two consecutive skids have different colours, and it requires a set-up process in the paint shop. One way of supplying the painting robots with paint is to lay a ring line up to the robot for each colour. Up to the paint spray gun, there is only one line in the robot. During a set-up process, this single line must be rinsed in order not to contaminate the new paint with the previous one. The time spent on this process is wasted and must be minimized. In addition, the colour set-ups always lead to residues of old colour and used solvents, which have to be disposed of at considerable expensive. The set-up times are always the same and do not depend on the colours. When a colour is set up, there is no limit to how many skids can be painted.

Painting sequence option one of Fig. 2 shows a possible sequence in which the number of set-ups is reduced in comparison to the second option. Potts and van Wassenhove [6] provided an additional notation for jobs with set-ups. The symbol s_f represents set-ups that occur sequentially and independently with each family of parts. In this case, a family is defined as a consecutive series of skids of the same colour. The first objective, the minimization of set-up processes or number of families, can be characterized as $\sum n_f$.

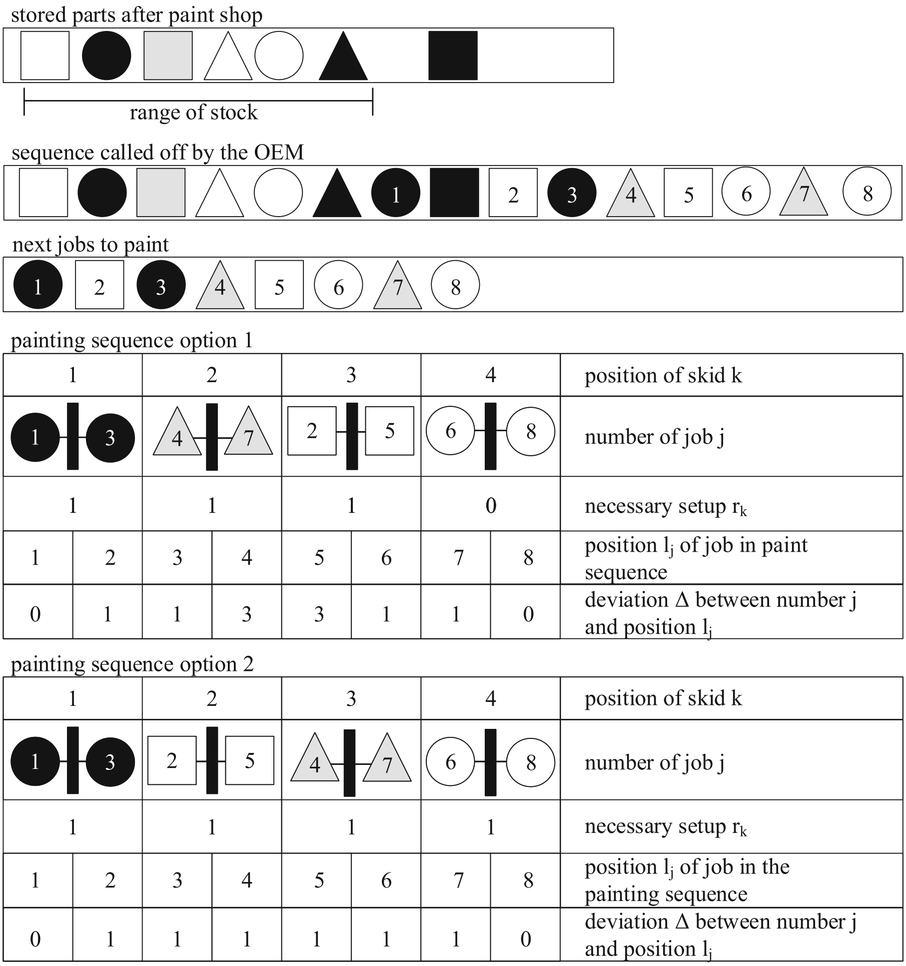


Fig. 2. Painting sequence options

As noted above, the second objective is to choose a painting sequence that is as close as possible to the sequence called off by the OEM except the already stored parts. The reason is to increase the stock range of painted parts which is determined by the number of parts with which the OEM sequence can be accommodated consecutively without a gap. As Fig. 2 shows, only six jobs can be included in the range of stock, although there are seven parts available, because the seventh order of the OEM sequence is missing. If, ideally, all stored parts match consecutively the sequence of the OEM, the highest level of safety would result in case of a malfunction in the paint shop. So the key is to identify a painting sequence including the missing jobs of the OEM sequence, and to minimize the deviations between the positions of each job in the painting and OEM sequence.

Strategically, a reduction in stocks can thus be achieved through this objective. If there are fewer gaps or single missing parts in the stock in comparison to the called-off sequence, the stock can be reduced with the same range of coverage. This means that long-term objectives can also be achieved by short-term sequencing, which was described by Pinedo and Chao [7]. The symbol Δ is introduced and represents the maximum of deviations between each job j of the OEM sequence and its assigned position l_j in the painting sequence. The objective is to minimize Δ and could also be expressed as the average deviation, the sum of all deviations, or the sum of all quadratic deviations. At this point, minimizing the maximum deviations is purposely decided because this value can be easily understood by everyone involved in a real-world application. Thus, the described problem can be characterized as $1|p_j = 1,\ s_f|\Delta,\ \sum n_f$, and this characterization can be used to compare and differentiate this problem from others.

Regarding to the characterization of T'kindt and Billaut [8], the problem can be defined as deterministic, repetitive and static. During the decision process, all jobs are known in advance, and no random influences are present. Due to constantly changing factors, it is necessary to plan a sequence several times a day. One of these factors is the rejection of painted components during inspection after painting. The sequence of approximately 40 skids is defined by a planning decision every hour, and approximately 1000 skids are planned a day in the focused real-world problem.

3 Related Literature

This article addresses a topic that deals with the components of various problems investigated in the literature:

- single-machine sequencing and scheduling,
- the due-date assignment problem, and
- the sequential-ordering problem.

Baker provided a definition for the difference between scheduling and sequencing. He pointed out that sequencing is a specialized case of scheduling in which the ordering of jobs determines a schedule [9]. The case of a single machine was first studied to minimize the maximum tardiness by Jackson [10]. Since then, many works on this subject have been published. An overview of deterministic *machine scheduling* can be found in the articles by Lawler et al. [1] and Anderson et al. [11]. Keha et al. [12] showed four different types of formulating single-machine problems using completion time variables, time index variables, linear ordering variables and assignment variables. The approach of using assignment variables is used in Sect. 4 for modelling the sequence in the proposed IP formulation.

To evaluate the performance of a schedule the following objectives are commonly used: schedule length, mean (weighted) flow time, maximum lateness, mean (weighted) tardiness, mean (weighted) earliness, and the number of early or tardy jobs and also combinations of them [13]. As early researchers, Seidmann et al. [14] addressed the combination of minimizing earliness and tardiness costs in the 1980s. The importance of JIT production increased at that time. The number of jobs completed after their due dates should be reduced in order to avoid paying penalty costs and/or losing customers. However, jobs completed early also result in disadvantages such as unnecessary inventory and insurance costs or the deterioration of goods [15]. Baker and Scudder [16] provided a review of this topic, which is also known as the *due-date assignment problem* or the earliness/tardiness (E/T) model.

In comparison to the problem considered in this paper, single-machine environments ($\alpha = 1$) are discussed, including set-ups ($\beta = s_f$). The minimization of earliness and tardiness of orders is based on due dates compared to the time of completion. This idea is also used in this paper, but can be seen as a simplified deviation Δ between the job position in the OEM sequence and the paint position, since all processing times are the same ($p_j = 1$), and the due date of delivery corresponds to the position in the OEM sequence. However, the minimization of set-ups takes place only as time or total costs in connection with the schedule variance. Other disadvantages of set-up processes are neglected, or weighting factors that are difficult to determine in practice have to be defined.

The minimization of set-up operations ($\sum n_f$) is explicitly considered in the *sequential-ordering problem*, which is used to minimize the changes of colour during the painting of car bodies in vehicle-manufacturing plants [17]. In particular, the possibility of changing a given sequence with the aid of mix-banks or other sorting systems is investigated [18]. This sequence can be changed only to a limited extent due to the capacity of the sorting systems. In the problem that is the focus of this paper, the changeability of the sequence is freely selectable. Furthermore, the deviation of jobs to the customer sequence is not considered. Taube and Minner [19] included the given sequence by the customer in their studies of a first-tier supplier, but the alteration of the sequence is also limited by physical constraints and release dates.

Although a number of papers consider several objectives or constraints of the focused problem, a model that can be applied to the described real-world situation has not been identified. Therefore, the next section presents an IP formulation.

4 IP Formulation

In the following an IP formulation is presented to solve the described problem, introduced as the *maximum deviation machine sequencing problem with set-ups* (MDMSP-S). The notations, shown in the Table 1 below, will be used.

Table 1. Notation for IP formulation

Indices	
j	Job
k	Skid
Decision variables	
x_{jk}	Binary assignment variable, 1 if job j is assigned to skid k
l_j	Position of job j in the sequence of assigned jobs
m_j	Position of the skid to which job j is assigned
Δ	Maximum deviation between each job j and its assigned position l_j
r_k	Binary set-up variable, 1 if a new colour is set up on skid k
c_k^{skid}	Colour assigned to skid k
t_k^{skid}	Type of skid k
y_k	Binary decision variable, 1 if the colour of skid k differs from that of skid $k-1$
Parameters	
N	Number of jobs that must be assigned to a skid
J	Number of given jobs
K	Number of given skids
c_j^{job}	Colour of job j
t_j^{job}	Type of job j
C	Number of different colours
T	Number of different types
M_1 to M_3	Large numbers

As explained above, there are two objectives to be taken into consideration. The first one, reducing colour set-ups, can be expressed as the pure number, the set-up time, or the monetary expenditure, whereby a conversion between the sizes is possible. The pure number is used in the following formulation as Eq. 1 shows. Therefore, the variable r_k, which assumes 1 if a colour change occurs on skid k, is summed up for all skids.

The second objective, preserving the OEM sequence, is to minimizing the deviation between the original position of a job j in the called-off sequence and the position l_j in which the order is painted. To avoid a real bi-objective formulation the *ϵ-constraint approach* is used. Therefore, the requirement of preserving the OEM sequence is modelled as an additional constraint in Eqs. 2 and 3. These inequalities ensure that the described deviations do not exceed Δ.

If both objectives were combined in one objective function, weightings, cost factors or a kind of transformation would be needed. In practice, however, it is difficult to determine these cost factors or weightings.

The ϵ-*constraint approach* uses a systematic variation of the additional constraints while optimizing the objective function [20]. In this case, Δ is increased

$$\text{Min.} \sum_{k=1}^{K} r_k \tag{1}$$

subject to

$$l_j - j \leq \Delta \qquad \forall\, j \tag{2}$$

$$j - l_j \leq \Delta \qquad \forall\, j \tag{3}$$

$$\sum_{k=1}^{K+1} x_{jk} = 1 \qquad \forall\, j \tag{4}$$

$$\sum_{j=1}^{J} x_{jk} = N \qquad \forall\, k \tag{5}$$

$$t_k^{skid} - t_j^{job} \leq M_1(1 - x_{jk}) \qquad \forall\, j, k \tag{6}$$

$$t_j^{job} - t_k^{skid} \leq M_1(1 - x_{jk}) \qquad \forall\, j, k \tag{7}$$

$$c_k^{skid} - c_j^{job} \leq M_2(1 - x_{jk}) \qquad \forall\, j, k \tag{8}$$

$$c_j^{job} - c_k^{skid} \leq M_2(1 - x_{jk}) \qquad \forall\, j, k \tag{9}$$

$$M_1 = T - 1 \tag{10}$$

$$M_2 = C - 1 \tag{11}$$

$$m_j = \sum_{k=1}^{K+1} x_{jk} \cdot k \qquad \forall\, j \tag{12}$$

$$l_j = \sum_{k=1}^{m_j - 1} \sum_{j^*=1}^{J} x_{j^*k} + \sum_{j^*=1}^{j} x_{j^* m_j} \qquad \forall\, j \tag{13}$$

$$c_k^{skid} \geq c_{k-1}^{skid} + 1 - M_3(1 - y_k^1) \qquad \forall\, k \tag{14}$$

$$c_{k-1}^{skid} \geq c_k^{skid} - M_3(y_k^1) \qquad \forall\, k \tag{15}$$

$$c_{k-1}^{skid} \geq c_k^{skid} + 1 - M_3(1 - y_k^2) \qquad \forall\, k \tag{16}$$

$$c_k^{skid} \geq c_{k-1}^{skid} - M_3(y_k^2) \qquad \forall\, k \tag{17}$$

$$r_k = y_k^1 + y_k^2 \qquad \forall\, k \tag{18}$$

$$c_0^{skid} = 0 \tag{19}$$

$$M_3 = C + 1 \tag{20}$$

$$x_{jk} = \begin{cases} 1 & \text{if job } j \text{ is assigned to skid } k \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

$$\Delta \in \{0, ..., J - 1\} \;\; j = 1, ..., J \quad l_j \in \{1, ..., J\} \tag{22}$$

$$k = 1, ..., K \quad m_j \in \{1, ..., K\} \tag{23}$$

$$y_k^1 \in \{0, 1\} \quad y_k^2 \in \{0, 1\} \quad r_k \in \{0, 1\} \quad M_1, M_2, M_3 \in \mathbb{N} \tag{24}$$

$$t_k^{skid} \in \{1, ..., T\} \quad c_k^{skid} \in \{1, ..., C\} \tag{25}$$

stepwise from a lower bound Δ_{min} to an upper bound Δ_{max}. Each variation leads to a pair of Δ and the minimized sum of r_k. A pair is called Pareto-optimal if it is not possible to improve the value of one objective without harming the other one [21]. The method for identifying the lower and upper bounds for Δ is described in Sect. 5.

A binary decision variable x_{jk} is introduced, which assumes the value 1 if job j is assigned to skid k. Equation 4 ensures that a job must be assigned to exactly one skid. It should be noted that an additional skid $K+1$ with unlimited capacity is introduced. This ensures that all orders that are not assigned to the skids $1...K$ can be allocated so that the solvability of the problem is enabled. This results from the real-world requirement that the capacity of the skids should be fully utilized to avoid empty slots. In the real-world case, there are approximately 1000 jobs known in advance that have to be distributed to approximately 40 skids during one planning decision, as previously mentioned in Sect. 2. Therefore, it is easily possible to fully load all skids.

The constraint 5 ensures that exactly N jobs are assigned to a skid k. The additional skid $K+1$ is not affected by this constraint.

To ensure that only parts or jobs of the same type are assigned to a skid, the constraints 6 and 7 are introduced. On the left side of the inequalities, the difference between the skid type t_k^{skid} and the job type t_j^{job} must be 0 if the job j is assigned to the skid k so that $x_{jk} = 1$. The *Big-M* formulation ensures that the constraints remain solvable if the order j is not assigned to the skid k so that $x_{jk} = 0$. The same procedure is used to ensure that only one colour can be painted per skid, as shown by the Eqs. 8 and 9. Since the difference between the largest and smallest value for a type or a colour is known, a value for M_1 and M_2 can be defined exactly, as Eqs. 10 and 11 show.

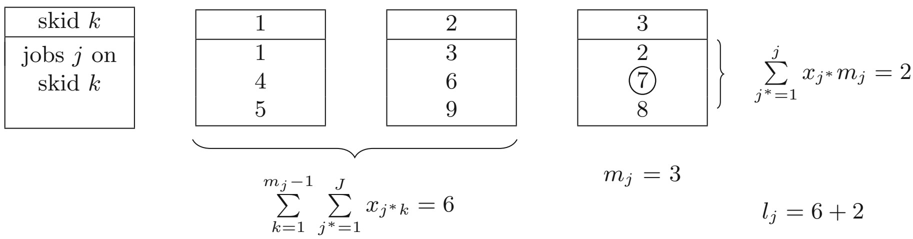


Fig. 3. Explanation: position of a job

Equation 12 is used to calculate the position of a skid m_j to which a job j is assigned. Figure 3 shows an example where $m_7 = 3$ is due to the assignment of job 7 to skid 3. The result will be used in the next constraint, 13, which defines the position of an order l_j in the painting sequence. The first summand of Eq. 13 counts all jobs that are assigned to the skids from 1 to $m_j - 1$. As Fig. 3 shows, the six orders of skid one and two are taken into account. The second summand adds the number of all orders on the skid, on which job j is assigned, if they have a job number smaller than j. In the example, the result is $l_j = 8$. An additional control variable j^* is required in the equation, since j is already blocked with the order for which the position is determined.

The constraints 14 to 18 determine whether a set-up occurs between two successive skids k and $k+1$. The variable c_k^{skid} describes the colour of the skid k based on the jobs assigned to it in Eqs. 8 and 9. To avoid logical expressions in the model, three cases must be distinguished. In the case $c_k^{skid} > c_{k-1}^{skid}$, the variable y_k^1 assumes the value 1 so that the inequalities 14 and 15 are valid. At the same time, variable y_k^2 assumes the value 0 in the constraints 16 and 17. For the second case $c_k^{skid} < c_{k-1}^{skid}$ the reverse applies: $y_k^2 = 0$, $y_k^2 = 1$. The third case represents the situation for which no set-up takes place. If c_k^{skid} and c_{k-1}^{skid} are equal, both variables y_k^1 and y_k^2 assume the value zero. Therefore the value 1 is added to the right side of the inequality in constraints 14 and 16, so that the binary variable assumes the value 1 if the colours are equal. Otherwise it could also be zero.

Equation 18 merges both binary variables. Thus, the variable r_k assumes the value 1 if a set-up operation is conducted before skid k. Constraint 19 is required to set an initial value to compare the colour with skid $k = 1$. Since c_0 is set to zero, the first skid always requires a set-up. In the real-world situation, the colour value would be taken from the last skid of the previous planning.

Table 2. Example of the constraints 14 to 18

c_k	$c_k - 1$	Equations 14–18	c_k and $c_k - 1$ inserted	y_k^1	y_k^2	r_k
5	0	$c_k^{skid} \geq c_{k-1}^{skid} + 1 - M_3(1 - y_k^1)$	$5 \geq 1 - 6(1 - y_k^1)$	1	–	1
		$c_{k-1}^{skid} \geq c_k^{skid} - M_3(y_k^1)$	$0 \geq 5 - 6(y_k^1)$			
		$c_{k-1}^{skid} \geq c_k^{skid} + 1 - M_3(1 - y_k^2)$	$0 \geq 6 - 6(1 - y_k^2)$	–	0	
		$c_k^{skid} \geq c_{k-1}^{skid} - M_3(y_k^2)$	$5 \geq 0 - 6(y_k^2)$			
2	3	$c_k^{skid} \geq c_{k-1}^{skid} + 1 - M_3(1 - y_k^1)$	$2 \geq 4 - 6(1 - y_k^1)$	0	–	1
		$c_{k-1}^{skid} \geq c_k^{skid} - M_3(y_k^1)$	$3 \geq 2 - 6(y_k^1)$			
		$c_{k-1}^{skid} \geq c_k^{skid} + 1 - M_3(1 - y_k^2)$	$3 \geq 3 - 6(1 - y_k^2)$	–	1	
		$c_k^{skid} \geq c_{k-1}^{skid} - M_3(y_k^2)$	$2 \geq 3 - 6(y_k^2)$			
4	4	$c_k^{skid} \geq c_{k-1}^{skid} + 1 - M_3(1 - y_k^1)$	$4 \geq 5 - 6(1 - y_k^1)$	0	–	0
		$c_{k-1}^{skid} \geq c_k^{skid} - M_3(y_k^1)$	$4 \geq 4 - 6(y_k^1)$			
		$c_{k-1}^{skid} \geq c_k^{skid} + 1 - M_3(1 - y_k^2)$	$4 \geq 5 - 6(1 - y_k^2)$	–	0	
		$c_k^{skid} \geq c_{k-1}^{skid} - M_3(y_k^2)$	$4 \geq 4 - 6(y_k^2)$			

Table 2 shows examples of the constraints 14 to 18 with three cases and five colours. The first one shows the maximum difference between the values of the colour, the second one a difference of 1, and in the third case, the colours are equal. Since the number of colours is five, M_3 can be set to 6, as Eq. 20 shows. The Eqs. 22 and 25 define the domain sets of the variables.

5 Illustrative Example

The following section provides with an example of the proposed model. There are 15 jobs available in three different colours and types, which should be assigned to 6 skids. The capacity of one skid is set to two jobs. Table 3 displays the input data.

Table 3. Input data

job j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
type of job t_j^{job}	3	2	2	1	3	2	3	1	1	2	2	3	1	3	3
colour of j c_j	1	1	3	2	2	1	1	1	2	1	3	2	1	2	1
Number of skids	5	Capacity of a skid								2					

As previously mentioned in Sect. 4 and described in the ϵ-constraint approach, it is necessary to find a lower and an upper bound for varying Δ. The upper bound can be easily defined as the maximum possible deviation that can occur: $\Delta_{max} = J - 1$. For choosing the lower bound, it must be noted that small values can lead to an unsolvable problem. In order to avoid this case, the problem should be solved first by minimizing Δ without considering the colour set-ups. The constraints 14 to 20 can be neglected as they have no influence on finding the smallest possible Δ. Afterwards, Δ is increased stepwise by 1 from Δ_{min} to Δ_{max} and for each Δ the minimum of r_k is calculated.

Table 4. Solution for $\Delta = 4$

skid k	1		2		3		4		5		6		7		
position l_j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
job j	2	6	1	7	4	9	3	11	5	12	8	13	10	14	15
type of job t_j^{job}	2	2	3	3	1	1	2	2	3	3	1	1	2	3	3
colour of job c_j	1	1	1	1	2	2	3	3	2	2	1	1	1	2	1
deviation $\|j - l_j\|$	1	4	2	3	1	3	4	3	4	2	3	1	3	0	0

The Table 4 contains the solution of the problem with $\Delta_{min} = 4$. Due to the minimal Δ, there are 5 colour changes necessary on the skids 1, 3, 4, 5 and 6. Therefore, the object function value assumes the value 5. It can also be seen that an additional skid, here skid 7, is introduced in the model. For this additional skid, the type and colour specifications as well as the capacity restrictions of the skids do not have any effect. It also has no impact on the objective function value. To obtain a solvable instance without this additional skid, the number of jobs per color type and job type would have to be exactly multiples of the

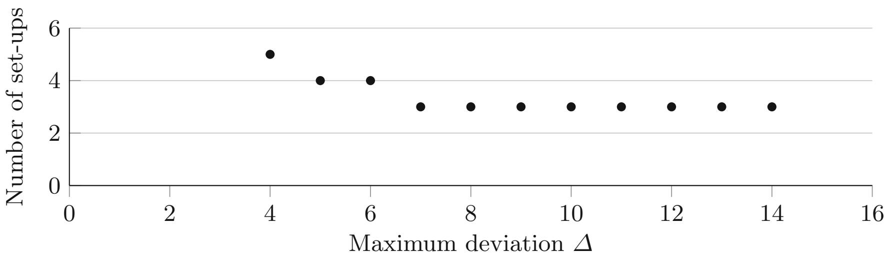


Fig. 4. Results of varying Δ

capacity N. For example, job 10 could only be assigned to one of the first six skids if there were another job with the same type and colour.

Figure 4 shows the corresponding minimum of $\sum r_k$ while increasing Δ incrementally by 1. The result is a set of pairs, which can be checked easily to determine if they are Pareto-optimal according to the definition given above. Those that are Pareto-optimal represent the solution to the problem.

Table 5. Solutions for $\Delta = 5$ and $\Delta = 7$

Maximum deviation $\Delta = 5$															
skid k	1		2		3		4		5		6		7		
position l_j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
job j	2	6	1	7	3	11	4	9	5	12	8	13	10	14	15
type of job t_j^{job}	2	2	3	3	2	2	1	1	3	3	1	1	2	3	3
colour of job c_j	1	1	1	1	3	3	2	2	2	2	1	1	1	2	1
deviation $\lvert j - l_j \rvert$	1	4	2	3	2	5	3	1	4	2	3	1	3	0	0
Maximum deviation $\Delta = 7$															
skid k	1		2		3		4		5		6		7		
position l_j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
job j	2	6	1	7	8	13	3	11	4	9	5	12	10	14	15
type of job t_j^{job}	2	2	3	3	1	1	2	2	1	1	3	3	2	3	3
colour of job c_j	1	1	1	1	1	1	3	3	2	2	2	2	1	2	1
deviation $\lvert j - l_j \rvert$	1	4	2	3	3	7	4	3	5	1	6	0	3	0	0

The pairs (4, 5), (5, 4) and (7, 3) are Pareto-optimal because no improvement in one objective function value is possible without worsening the other. Table 5 shows the solution of the Pareto-optimal pairs (5, 4) and (7, 3).

In the case of $\Delta = 5$, it can be seen that order 11 is assigned earlier, which saves a set-up between colours 2 and 3. In the case of $\Delta = 7$, it is possible to assign orders 8 and 13 already on skid 3, which is why no further change to colour 1 is required in the following sequence.

Finally, a decision must be made regarding which Pareto-optimal solution is to be used to define the sequence for lacquering. This should depend on the stock level. If the stock level is low, a solution with a low maximum deviation must be chosen to avoid more frequent gaps in the OEM sequence. In this case, the costs and disadvantages of the more-frequent set-ups have to be accepted. If the stock level is high, the reverse case occurs.

At this point, it should be noted that several solutions for a given deviation may arise. In the example of $\Delta = 5$, jobs 5 and 12 on skid 4 can be exchanged with jobs 4 and 9 on skid 5, and the solution will remain valid. This represents a deterioration in the sense of the stated objective. It is possible to avoid this problem within postprocessing. For this purpose, the orders assigned on consecutive skids of the same colour must be sorted in ascending order by the sum of their job numbers. To avoid violating the constraints of job types, only all orders of a skid can be exchanged.

6 Computational Results

The proposed model was implemented and tested for six instances on an *Intel(R) Core i7(R) 3537U CPU* with 3.1 GHz clock speed and 12 GB RAM using *CPLEX Optimization Studio 12.8* with max. 4 parallel threads. Each instance has a given number of jobs and skids and a defined capacity for each skid. The job types and colours were created randomly according to a uniform distribution and the run time was limited to 60 min. The objective is to find the smallest possible Δ for each instance and its corresponding sum of r_k. The results are shown in Table 6.

Table 6. Computational results

Number of jobs J	Number of skids K	Capacity of the skid N	Number of colours C	Time to first IP solution [s]	Time to last IP solution [s]	GAP of last IP solution	Proof of optimality
30	5	2	2	3.4	4.6	50%	yes
30	10	2	2	3,8	64	10 %	yes
40	10	2	2	4.7	110	20 %	yes
40	16	2	2	25,5	481	10 %	yes
50	20	2	2	18	1291	10 %	yes
150	30	3	10	345	3600	46,83%	no

It can be seen that the first five instances were solved to optimality, but only the first four instances were solved in an acceptable runtime. Runtimes under 10 min are acceptable for use in real-world situations. As mentioned in Sect. 2, there are a number of influences that make it necessary to generate the next painting sequence every hour. That means it is not acceptable to plan a sequence with data that is older than 10 min. The instances with 30 and 40 jobs show that runtimes depend not only on the number of jobs themselves but also on the input data, especially the number of skids, which has a significant influence.

Instance number 6 with 150 jobs shows an almost real-world scenario. Due to the higher capacity of skids and the number of colours, the complexity increases. The GAP of the first IP solution is 48.46%. With a time limit of 60 min, the result could be improved only to a GAP of 46.83%. It can be stated that the runtimes are strictly not applicable in real-world situations. Therefore, further research on more-efficient modelling of the proposed problem is desired and the application of heuristics should be considered in order to achieve acceptable runtimes.

7 Conclusion

In this article, a real-world problem was presented in order to solve the decision-making problem in the formation of sequences in the paint shop of a first-tier supplier in the automotive industry. For this purpose, the situation has been described in detail, and a characterization of the problem has been provided. Several articles in the literature showed similarities with the described problem in several aspects, but none of the studies could represent the described real-world case. Therefore, a model has been presented, and the procedure for the solution has been described. In addition, computational studies were provided showing that the runtimes are not acceptable.

In future investigations, the complexity of the problem should be determined in order to identify more-efficient solution methods. One possibility could be using a heuristic to find satisfactory solutions with acceptable runtimes.

Several of the presented properties of the problem are realized in technically different ways by other suppliers. A paint-supply system with cartridges that can be filled with only a certain amount of paint can be used. Furthermore, not all orders or skids can be available immediately, or one skid can be loaded with different types of parts. These examples would influence the constraints, and by taking them into consideration, the model could be used in more cases.

In order to increase the solution quality, modified formulation of the objective function should be investigated in order to avoid the problem described in Sect. 5. Furthermore, the criticality of the orders could be given more consideration. The painting processes of plastic parts also have certain reject rates. These stochastic influences should also be analysed. The mentioned aspects can be addressed by further research.

References

1. Lawler, E.L., Lenstra, J.K., Kan, A.H.R., Shmoys, D.B.: Sequencing and scheduling: algorithms and complexity. In: Logistics of Production and Inventory. Handbooks in Operations Research and Management Science, vol. 4, pp. 445–522. Elsevier (1993). Chapter 9

2. Wilbrecht, J.K., Prescott, W.B.: The influence of setup time on job shop performance. Manage. Sci. **16**(4), 274–280 (1969)
3. Boysen, N., Scholl, A., Wopperer, N.: Resequencing of mixed-model assembly lines: survey and research agenda. Eur. J. Oper. Res. **216**(3), 594–604 (2012)
4. Meyr, H.: Supply chain planning in the German automotive industry. OR Spectrum Quantit. Approaches Manag. **26**(4), 447–470 (2004)
5. Graham, R., Lawler, E., Lenstra, J., Rinnooy Kan, A.: Optimization and approximation in deterministic sequencing and scheduling: a survey. In: Hammer, P.L., Johnson, E.L., Korte, B.H. (eds.) Discrete Optimization II. Annals of Discrete Mathematics, vol. 5, pp. 287–326. Elsevier (1979)
6. Potts, C.N., Van Wassenhove, L.N.: Integrating scheduling with batching and lotsizing: a review of algorithms and complexity. J. Oper. Res. Soc. **43**(5), 395–406 (1992)
7. Pinedo, M., Chao, X.: Operations Scheduling: With Applications in Manufacturing and Services. Irwin/McGraw-Hill, Boston (1999)
8. T'kindt, V., Billaut, J.-C.: Multicriteria Scheduling: Theory, Models and Algorithms. Springer, Heidelberg (2006)
9. Baker, K.R., Trietsch, D.: Principles of Sequencing and Scheduling. Wiley, Hoboken (2009)
10. Jackson, J.R.: Scheduling a production line to minimize maximum tardiness. Research report. Office of Technical Services (1955)
11. Anderson, E.J., Glass, C.A., Potts, C.N.: Machine scheduling. Local Search Comb. Optim. **11**, 361–414 (1997)
12. Keha, A.B., Khowala, K., Fowler, J.W.: Mixed integer programming formulations for single machine scheduling problems. Comput. Ind. Eng. **56**(1), 357–367 (2009)
13. Błażewicz, J., Ecker, K.H., Pesch, E., Schmidt, G., Węglarz, J.: Handbook on Scheduling: From Theory to Applications. Springer, Heidelberg (2007)
14. Seidmann, A., Panwalkar, S.S., Smith, M.L.: Optimal assignment of due-dates for a single processor scheduling problem. Int. J. Prod. Res. **19**(4), 393–399 (1981)
15. Shabtay, D.: Scheduling and due date assignment to minimize earliness, tardiness, holding, due date assignment and batch delivery costs. Int. J. Prod. Econ. **123**(1), 235–242 (2010)
16. Baker, K.R., Scudder, G.D.: Sequencing with earliness and tardiness penalties: a review. Oper. Res. **38**(1), 22–36 (1990)
17. Lin, L., Sun, H., Xu, Y.-q.: A research of resequencing problem in automobile paint shops using selectivity banks. In: IEEE 18th International Conference on Industrial Engineering and Engineering Management (IE&EM), pp. 658–662 (2011)
18. Spieckermann, S., Gutenschwager, K., Voß, S.: A sequential ordering problem in automotive paint shops. Int. J. Prod. Res. **42**(9), 1865–1878 (2004)
19. Taube, F., Minner, S.: Resequencing mixed-model assembly lines with restoration to customer orders. Omega Int. J. Manag. Sci. **78**, 99–111 (2018)
20. Marler, R., Arora, J.: Survey of multi-objective optimization methods for engineering. Struct. Multidisc. Optim. **26**(6), 369–395 (2004)
21. Pinedo, M.: Scheduling. Springer (2012)

Decomposition Strategies for Multi-network Crew Scheduling with Attendance Rates for Conductors

Julia Heil(✉)

TU Dresden, 01062 Dresden, Germany
jheil@tu-dresden.de

Abstract. Railway crew scheduling is the problem of generating feasible duties for the crews on a train to cover all trips at minimal cost. In Germany, regional passenger transportation consists of many distinct but interlinked networks, each with own crews. For efficiency reasons, we investigate the cost saving potential of scheduling crews collectively across multiple networks operated by the same company. To derive valid estimates, we develop a solution approach for the large-scale multi-network crew scheduling problem considering the network-specific constraints of attendance rates for conductors. Several studies have shown that partitioning large-scale problems improves computational times. We discuss guidelines for a problem-specific decomposition and derive three methods: a graph partitioning algorithm with adjusted edge weights and two variants of a network-based greedy decomposition heuristic. We assess their performance with a 2-phase optimization method using a hybrid column generation genetic algorithm and benchmark the results against a test run without decomposition. The tests show that maintaining the network structure while considering the connectivity between networks achieves the best results.

1 Introduction

Crew scheduling is an important step of the operational planning in railways. It deals with generating duties to cover all trips at minimal cost while considering numerous real-world requirements. This work deals with the problem of scheduling conductors for multiple regional railway networks with the network-specific constraints of attendance rates, a problem arising at a large railway company.

Railway passenger transportation on regional level connects smaller cities and towns with each other and, if existing, with larger cities in the same area. It serves the regional and local travel demand, e.g., for commuting, leisure activities, or connection to larger cities and other transportation modes. The train lines are typically designed for stopping at almost every station within the range of a certain distance. In Germany, local state transport authorities plan these railway networks for each geographical area. These separately defined networks are interlinked to ensure travel connections as can be seen from Fig. 1.

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 298–313, 2019.
https://doi.org/10.1007/978-3-030-29821-0_20

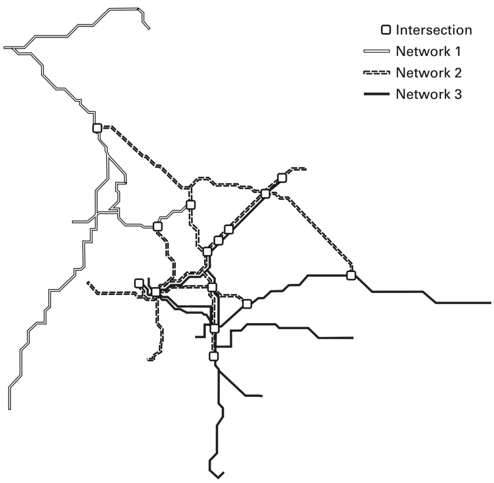


Fig. 1. Geographical extend of three regional passenger transportation networks with 15 important intersections

After defining a network by its train lines and corresponding timetable, the authorities select an railway company by means of a public tender. The transportation contract also formalizes service requirements such as the type and quality of the rolling stock, the ticket pricing structure or attendance rates for conductors. The latter is a special characteristic in the German regional railway passenger transportation and has to be considered when scheduling conductors to operate the network. Conductors' work mainly consists of passenger services (e.g., controlling tickets) and sometimes of operational tasks (e.g., securing departures). Hence, in contrast to drivers, they are not required on a train during the total driving time. Mainly due to cost reasons, the state transport authorities assign each train kilometer an attendance rate g which may vary depending on product type, time of day and other factors. Attendance rate g is satisfied when the g-share of all kilometers assigned to g is covered by at least one conductor. It can be defined as

$$g \leq \frac{\text{train kilometers}_{g,\text{attended}}}{\text{train kilometers}_{g,\text{total}}}. \tag{1}$$

Based on the terms of the transportation contract and further operational and legal restrictions, the nominated railway company schedules its resources, i.e. rolling stock and crews, to operate the network. The cost for rolling stock and personnel are two main cost factors of railway operations. Since companies strive for resource-efficient schedules, they commonly support these scheduling tasks with software based on Operations Research techniques. Due to the public tender process, companies plan each single network separately, even though schedules that operate across multiple interlinked networks could achieve economies of scale. Only in few cases, planners use their knowledge and experience to create duties across single networks manually. In our project, we investigate the cost saving potential of planning crews across multiple networks with attendance rates, a problem arising at the largest German railway operator.

Solving single networks with attendance rates has been extensively discussed by [8,9]. Our solution approach for solving multiple networks with attendance rates builds on the hybrid column generation genetic algorithm developed by [8]. In a previous study, we have shown that partitioning the original multi-network problem into smaller sub-problems enables parallel computation and thus acceleration of the solution procedure. By sub-sequentially solving the original problem building on the solutions of the previous phase, the algorithm regains solution quality. With this approach, cost savings up to 2% for 12 networks with a planning horizon of one day could be realized [6]. In this study, each single network represented a sub-problem which led to two drawbacks: Firstly, the needed CPU capacity for parallel computation increases proportionally with the number of sub-problems to be solved. Secondly, the different sizes and complexities of the sub-problems led to a high variance in computational times in the partitioning phase. We aim for a limited use of CPU capacity and balanced computational times by creating few sub-problems of similar size and complexity. To achieve this, different decomposition strategies suitable for the multi-network crew scheduling problem with attendance rates are investigated.

In this work, we describe the multiple network crew scheduling problem with attendance rates and its size and complexity (Sect. 2). We discuss related work (Sect. 3) and present an improved 2-phase optimization method based on a hybrid column generation genetic algorithm (Sect. 4). In Sect. 5, we outline problem-specific considerations to decompose the original problem which translate into three different decomposition approaches. We assess their performance with two real-world instances and benchmark the results against the solution method without decomposition (Sect. 6). We conclude with a summary of our findings and potential for further research (Sect. 7).

2 Problem Description

Crew scheduling is part of the planning process for tenders as well as the operational planning at a railway company. Its objective is to select a set of feasible duties, a *schedule*, which covers all trips at minimum cost. A *trip* represents an atomic unit of a train run and is characterized by a departure/arrival time and a departure/arrival station (*relief points*). Relief points are stations where crew members can change trains. Some of them are equipped for taking breaks. Trips i and j can be combined into feasible *duties*, if

1. the starting relief point of trip j is the end relief point of trip i (*geographical condition*) and
2. trip j starts after trip i with a maximum waiting time (*chronological condition*).

If the trips i and j are not on the same train run, the walking distance from one train to the other must be considered. We set the maximum waiting time and the walking time to 120 and 5 min, respectively. Also, duties can only start and end at a few defined stations (*crew depots*) which have to be identical (*duty symmetry*). Further restrictions include additional operational conditions (e.g., *break opportunities* at relief points) and legal work regulations (e.g., *maximum working time*, *break time rules*). A comprehensive discussion of all restrictions that apply to our real-world problem for the single network case can be found in [8,9].

2.1 Attendance Rates in the Multi-network Railway Crew Scheduling Problem

For being able to schedule crews across multiple networks, it is necessary to combine separately defined networks and their specific restrictions into a crew scheduling problem (*multi-network railway crew scheduling problem*).

We define a *single railway network* as the entity of train lines and the corresponding parameters determined by a transportation contract. Obviously, it is only reasonable to combine two single networks if they are interlinked, i.e. the two networks share at least one relief point (an *intersection* of two or more networks). Otherwise, crews cannot change between the trains of different networks. Combining two or more single networks results in a very large synthetic network of trips and relief points with their corresponding characteristics and restrictions.

Not all restrictions apply to every network. Attendance rates for conductors, in particular, are defined for each network individually. However, other restrictions and characteristics such as the aforementioned operational conditions or legal work regulations usually apply to a station (e.g., walking distances), to a duty (e.g., maximum working time, break time rules, symmetry of starting/end crew depot) or to the schedule in total (e.g., average working time). Hence, we define the problem in such a way that attendance rates are the only network-specific restrictions, i.e. must be satisfied for the set of trips of each network individually.

2.2 Mathematical Formulation

We formulate the *multi-network railway crew scheduling problem with attendance rates (MNRCSPAR)* as set covering model. Let M and N denote the set of trips i and the set of duties k, respectively. Furthermore, we define R as the set of networks r and G_r as the set of required attendance rates g per network r. The binary decision variable x_k indicates whether duty k is selected in the solution schedule, i.e. $x_k = 1$, 0 otherwise. Let y_i be a binary decision variable such that $y_i = 1$ if trip i is covered by the schedule, 0 otherwise. We define the travel distance of trip i which is assigned to attendance rate g as d_{ig} and the cost of duty k as c_k. The binary assignment matrix A indicates whether trip i is covered by duty k ($a_{ik} = 1$) or not ($a_{ik} = 0$). The model formulation is:

$$\min \sum_{k \in N} c_k x_k \tag{2}$$

$$s.t. \sum_{i \in M_r} d_{ig} y_i \geq g \sum_{i \in M_r} d_{ig} \quad \forall g \in G_r, \forall r \in R \tag{3}$$

$$\sum_{k \in N} a_{ik} x_k \geq y_i \qquad \forall i \in M \tag{4}$$

$$y_i \geq a_{ik} x_k \qquad \forall i \in M, \forall k \in N \tag{5}$$

$$x_k, y_i \in \{0, 1\} \qquad \forall i \in M, \forall k \in N. \tag{6}$$

The objective (2) is to minimize the total cost of all duties. Constraints (3) ensure that the total distance of attended trips of network r with attendance rate g satisfies the by attendance rate g defined minimum (see Eq. (1) in Sect. 1). Constraints (4) and (5) link the trip coverage by the duties of the solution schedule to the attendance rate constraints (cf. constraints (3)): if the minimal attended distance for attendance rate g requires that trip i is covered, then at least one duty k has to cover trip i. Vice-versa, constraints (5) ensure that if trip i is covered by at least one duty of the solution schedule, its distance d_i adds to the corresponding attendance rate fulfillment.

2.3 Size and Complexity of the MNRCSPAR

A MNRCSPAR is more complex than single network problems for two reasons. First, the total number of trips to be covered is much larger since it consists of the sum of all trips of each network plus additional trips. These are created when a network serves a station which is not defined as relief point, but is a relief point in another network. As a consequence, some trips are broken up into two or more trips as compared to the single network case. This adds both restrictions and variables, i.e. duties to cover all trips, to the problem.

Secondly, the total number of possible trip combinations, i.e. duties, increases with the higher number of possible successors per trip. Figure 2 illustrates the typical distribution of possible successors per trip for different network types.

Typical small to medium-size networks are composed of a low to medium number of lines with a moderate frequency (e.g., every 60 min). Most trips can be combined with 1 to 10 successors, the maximum ranges from 10 to 15. In contrast, large and complex networks are characterized by a high number of train lines and/or a high frequency of train services with a short average trip duration. In our example, the distribution of the number of possible trip successors shows the highest peak at 38. Those are trips which end at one of the most important junction of the network where several train lines interconnect. The service frequency of these lines is relatively high, every 20 or 30 min, and remains constant throughout the day. Additionally, the frequencies of the train lines are synchronized in such a way that passengers have enough time to change

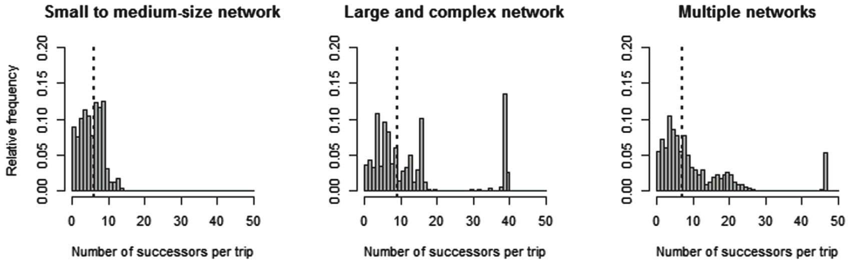


Fig. 2. Typical distribution of the number of possible successors per trip for a small to medium-size network, a large and complex network and multiple networks

trains. Hence, each trip which ends at this junction shows the same pattern of possible successors throughout the day.

If both of such network types are combined to multiple networks, the maximum number of possible trip successors increases, in this example up to 48. The median is slightly reduced as compared to the large and complex single network, but is higher than for the typical small to medium-size network. The average trip duration in our real-case data set is about 18 min. As a consequence, duties consist of 15 to 20 trips on average, in extreme cases up to 40. With the higher number of potential successors, the possibilities for combining trips to feasible duties increases to an even larger extend resulting an immense number of variables in the MNRCSPAR.

Crew scheduling problems are known to be NP-hard and therefore require high computational times for optimal solutions. This is not suitable for the planning tasks of our railway company. Its planners generate crew schedules for tender offers and for the daily operations. Planning for tenders commonly occurs every 10 years per network, depending on the duration of the transportation contract, and is a complex task including many dependencies with the previous planning steps. Typically, it takes several months until a tender offer is ready for submission. Operational planning occurs more frequently but is a little less extensive as it can build on the knowledge and experience of the previous schedules. It is common to review and adjust the crew schedule for each network on a yearly or half-yearly basis, following the changes of the time tables. However, due to potential short-term changes (e.g., construction sites), intrayear adjustments to the schedule might become necessary. For both planning tasks, the planners test different schedule parameters and modify the data set in order to improve the overall cost in a number of iterations. Hence, to support the planning tasks, our software tool should generate feasible and high quality solutions which achieve lower cost than the sum of the solutions of the single network problems within less than a few days. These requirements necessitate a tailored method to compute high quality results for the multi network railway crew scheduling problem in reasonable time.

3 Related Work

Researchers have dealt with crew scheduling problems in the transportation industry since the 1950s. Originally, their focus was on problems in the public transit and airlines industries which still are areas of high interest (see, e.g., the recent surveys of [10,15]). In the 1990s, due to the advances of computational power, research extended to the more complex problems of crew scheduling in railway systems (e.g., see [2,7]). Since then, a large number of real-world applications have been studied varying in, among others, transportation modes, countries and their specific requirements, as well as the integration level of planning stages.

Heuristic procedures and meta-heuristics are frequently employed to achieve faster computational times. By far the most popular solution method is column generation (see [18] for a methodological overview) with various forms and modifications to fit the specific real-world case. In the context of crew scheduling, the problem of selecting duties for the best schedule starts with a small subset of feasible duties (*restricted master problem*). It is stepwise enlarged by newly generated duties, i.e. columns, which potentially improve the current solution. Generating these duties is an optimization problem by itself (*pricing problem*). Since the master problem is relaxed by linear or Lagrangean relaxation, integer solutions are either generated by a branch-and-bound scheme (*branch-and-price*) or in a subsequent step.

Due to its slow convergence, researchers investigate problem-specific techniques to achieve better solutions in less time. In general, we distinguish between *acceleration techniques* for the solution method itself, *problem size management by merging and restrictions* and *problem size management by decomposition*. For common *acceleration techniques* such as column deletion and column fixation, [3] provide an comprehensive overview.

Problem size management by merging and restrictions includes approaches to reduce the graph size of the original problem by dissolving defined parts of it or to reduce the solution space by adding limits. [17] discuss three main ways of controlling the search space to extend their column generation approach by iteratively deriving new sub-problems of manageable size for the commercial solver (to continuously improve the current solution). These search space controllers include, among others, dissolving relief opportunities using experienced-based rules and limiting the number of trips per duty. Additional restrictions such as a minimal threshold for the share of working time or minimal duty duration reduce the solution space further. Likewise, [5] add limits to the number of possible connections per trip.

Problem size management by decomposition uses strategies to partition the graph of the original problem into a number of smaller sub-graphs which are solved either independently or remain linked to each other. Logical dimensions to decompose a problem are time, geography, train lines and historical schedule information (see [13]). [11] leverage the periodic element of a weekly RCSP to transform it into smaller one-period problems using the concept of time frames. Geographical and train line aspects are used by [16] who limit the trips for a problem instance to certain railway lines or train types or by a maximum distance from a depot. Likewise, [19] formulate individual problems per geographical crew district in order to cope with the large problem size. [4] develop two overlapping sub-problems based on the geographical position of the origin and destination of the trips. [1] experiment with all of the above mentioned partitioning dimensions: weekday partitioning, geographical partitioning, line based partitioning and partitioning based on information on good trip combinations from previous optimization runs. They show that combining the partitioning methods with the original solution procedure achieves a solution improvement by up to 2%. Some researchers investigate the use of algorithms to decompose the problem. [20] use a genetic algorithm to assign trips to depots and solve the smaller instances independently. Extensive research to partition graphs for problems at a freight railway operator was done by [12,13]. They explore and compare different edge weights for a graph partitioning problem and solve the resulting sub-problems dynamically using a divide-and-price algorithm.

With such problem size management approaches, computational times can be reduced significantly, but with a trade-off between reduction of solution space and loss of quality (see [17]).

4 Solution Approach

We build on the existing approach for the single network multi-day crew scheduling problem with attendance rates: a hybrid column generation genetic algorithm referred to as CGGA in the following (see [8] for a detailed discussion). In order to accelerate computational time for the MNRCSPAR, we experiment with the decomposition of the original problem into a number of sub-problems. Besides reducing the problem size, this enables parallel computation. However, as we aim for a better solution quality as the sum of the single network problems, it seems not sufficient to only decompose the original problem into smaller sub-problems and add up their solutions in the end. Instead, we expect to generate more cost savings by solving the complete original problem in a later phase. As a consequence, we suggest a 2-phase optimization method (see Fig. 3).

Phase 1 starts with the decomposition of the original problem into smaller sub-problems which are solved in parallel thereafter. We discuss different decomposition strategies and algorithms in Sect. 5. The sub-problems as well as the original problem in the later phase are solved using CGGA modified for network-specific constraints. Within CGGA, the linear relaxation at the root node of the master problem is solved with column generation. It utilizes a genetic algorithm for generating new duties, i.e. columns. The initial set of duties consists of efficient trip combinations which are built by three different graph search strategies. The column generation procedure applies column deletion to manage the size of the master problem and ends as soon as no new column with negative reduced cost could be generated by the genetic algorithm. Other stopping criteria such as time limits or limiting the number of iterations can also be applied. Instead of computing an integer solution for each sub-problem, we select every column

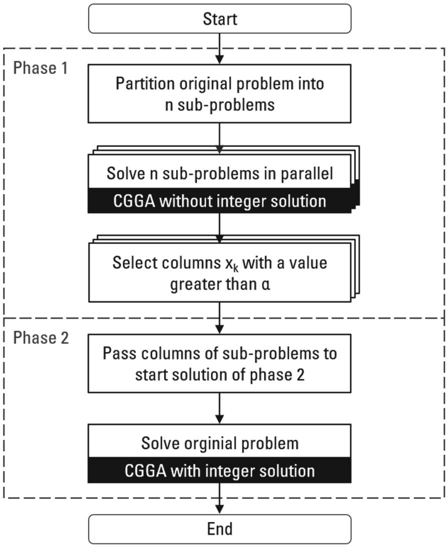


Fig. 3. Framework of the 2-phase optimization method with variable selection between phase 1 and 2

x_k with a primal value greater than a defined threshold α and pass it on to the second phase. Subsequently, in phase 2, we solve the original problem using CGGA building on an initial set of duties which combines the selected duties of each sub-problem and additional efficient trip combinations. In the last step, the integer solution is generated using a commercial solver.

Note that the initial set of duties of the second phase does not necessarily represent a feasible solution. The quality of the duty set per sub-problem highly depends on the decomposition which affects the sub-problems structure and its solvability, in particular with regard to the attendance rates requirements. Hence, suitable decomposition strategies are required.

5 Decomposition Strategies

An essential characteristic of the MNRCSPAR is that some restrictions, i.e. attendance rates, only apply to a subset of trips. Maintaining this structure of restrictions might support the solvability of the sub-problems. Additionally, to build a powerful initial set of duties to be further improved by the genetic algorithm of CGGA in phase 2, it seems beneficial to combine two sets of duties: duties which represent good solutions of the sub-problems considering attendance rate restrictions and duties which are efficient trip combination across all networks without considering attendance rate restrictions. The latter are built from the in the first step of CGGA, while the first must be derived from the sub-problem solutions using a suitable decomposition strategy. To test this assumption, we compare the performance of different strategies with the 2-phase optimization method.

5.1 Guidelines for Decomposing the Graph of a MNRCSPAR

Our objective is to decompose the MNRCSPAR into a defined number of sub-problems by cutting as few "good" trip combinations as possible.

The underlying structure of the MNRCSPAR can be interpreted as directed, weighted graph $G = (V, E)$ with vertices $v \in V$ representing trips. An edge $(i, j) \in E \subset V \times V$ with edge weight w_{ij} exists, if two trips i and j can be connected in accordance with the geographical and chronological conditions and the maximum waiting time (see Sect. 2). Finding a good decomposition is a graph partitioning problem [1], [13]. It is formally described for a graph G to find a partition of V into k equal-sized subsets $V_1, ..., V_k$ with minimum edge cut w_p, where

$$w_p = \sum_{n=1}^{k} (\sum_{(i,j) \in E : i \in V_n, j \in V \setminus V_n} w_{ij} + \sum_{(i,j) \in E : i \in V \setminus V_n, j \in V_n} w_{ij}) \quad (7)$$

In our case, the importance of an edge, represented by its weight, is determined by three factors (see [13] for a detailed discussion of the first two items):

- **Cost efficiency** - Unproductive times, e.g., crews waiting for the next trip, increase the cost of an schedule and, therefore, should be minimized.
- **Exclusiveness of a connection** - With an increasing number of outgoing edges of a vertex, the importance of one of these edges decreases as other alternatives exist. However, if a connection between two trips is unique for one of them, it should not been cut. Otherwise, the trip can only become the first or last trip of a duty which might lead to an insolvable problem instance.
- **Network or train run affiliation** - Trips of the same network or train run should be kept together in one sub-problem to reflect the problem structure with attendance rates.

Furthermore, the decomposition method applied should generate an adjustable number of sub-problem, e.g., ranging from 2 to 8 or more. The reasonable number of sub-problems depends on the problem size, its network structure and the available computational power for solving the MNRCSPAR.

5.2 Decomposition by a Graph Partitioning Algorithm

The NP-hard graph partitioning problem has been studied extensively. Various heuristics have been proposed to generate good solutions in very short time. In the context of railway crew scheduling problems, [13] show that adjusted edge weights and subset size modifications improve the performance of their solution method.

For our purpose, we define problem-specific edge weights (cf. Eq. 8) to represent cost efficiency, exclusiveness of a connection and train run affiliation. Cost efficiency is driven by the unproductive time, i.e. the waiting time t_{ij} in minutes, of a connection. We use the inverse of the waiting time as indicator (cf. first addend of 8). The exclusiveness of an edge can be measured by the inverse of the number of alternative edges. An edge e_{ij} is always both one of the outgoing edges (e_{iv}) of vertex i as well as one of the incoming edges (e_{vj}) of vertex j. It will therefore be assigned two values of which the maximum is chosen (cf. second addend of 8). Lastly, if the two vertices i and j belong to the same train run, indicated by the train number tNo, the weight of their linking edge is increased by a constant $z_{ij} = 1$ (cf. third addend of 8). With this, the addends are defined in such way that all three factors are considered equally with a value range of $(0, 1]$ and $[0, 1]$, respectively.

$$w_{ij} = \frac{1}{t_{ij}} + \max\left\{\frac{1}{\sum_{v \in V} e_{vj}}, \frac{1}{\sum_{v \in V} e_{iv}}\right\} + z_{ij} \text{ with } z_{ij} = \begin{cases} 1 & \text{if } tNo_i = tNo_j \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

We use the algorithm developed by [14], a multilevel recursive-bisection which has proven to achieve good and fast results. The algorithm and further improvements are implemented in the open-source software tool METIS[1], version 5.1. It enables k-partitioning of an undirected, weighted graph into k subsets which fits our purpose.

5.3 Decomposition by Network-Based Greedy Heuristics

As we assume that maintaining the single network structure supports our solution procedure, we also define a network-based greedy heuristic for decomposing the MNRCSPAR. Next to maintaining the single network structure, we aim for a balanced number of trips across the sub-problems in order to achieve comparable computational time and effort. The basic idea is that, given the trips of the original problem and a target number of sub-problems, a set of trips of a network is selected according to a defined rule and added to a sub-problem. The number of trips per sub-problem is hereby limited to 110% of the average number of trips per sub-problem.

Variant 1. First, we order the single networks by their count of trips. Then, starting with the network with the highest count of trips, the sets of trips per network are iteratively added to the sub-problem with the currently lowest number of trips. If the limit for trips is exceeded after adding a set of trips to a sub-problem, the sub-problem is "closed" and cannot be considered any further. The procedure continues until all sets of trips, i.e. single networks, are assigned.

This variant considers only the number of trips in order to achieve balanced sub-problem sizes. However, the degree of connectivity between single networks

[1] METIS is a set of serial programs for, among others, partitioning graphs and partitioning finite element meshes. See http://glaros.dtc.umn.edu/gkhome/metis/metis/overviewformoreinformation.

can vary significantly. For instance, if the MCRCSPAR includes more than two networks, it is possible that some of them are only indirectly connected via other networks. i.e. they do not share any edge. Since phase 1 of the solution procedure should already generate feasible duties which potentially are part of the final solution, considering connectivity between networks seems to be useful.

Variant 2. In the first step, we calculate the connectivity between all single networks defined as the number of edges connecting the sets of trips of two networks. The connectivity values between each network and each sub-problem are set to 0. We start with the highest connectivity value and add the trips of both networks r_1 and r_2 to the same sub-problem sp_1. It follows the update of the connectivity values in such a way that the connectivity between a network r_3 and sub-problem sp_1 is updated to the sum of the connectivity values between network r_3 and the networks r_1 and r_2, respectively. The latter are set to 0 afterwards. We iterate this procedure of selecting the highest connectivity value, assigning the trips to a sub-problem and updating the connectivity values until all sets of trips are assigned. As soon as the limit number of trips is exceeded after adding a set of trips to a sub-problem, all connectivity values of the corresponding sub-problem are set to 0 ("closing" of sub-problem). If a network remains with connectivity values of 0 to all of the open sub-problems, its trips are added to the sub-problem with the lowest number of trips.

6 Computational Results

We apply the decomposition strategies (see Sect. 5) to the partitioning step in phase 1 of the 2-phase optimization method (2PH, see Sect. 4). As a result, three alternative solution methods are derived: 2PH in combination with the graph partitioning algorithm (2PH_GP) and with the two variants of the network-based greedy decomposition heuristic (2PH_H1 and 2PH_H2). We evaluate the performance of the three alternatives with two real-world test instances (see Table 1).

Table 1. Structure and size of two real-world test instances

	Single networks [#]	Planning period [days]	Total trips [#]	Relief points [#]	Crew depots [#]
Test instance 1	2	7	18.286	51	10
Test instance 2	11	1	6.364	136	29

All tests are executed on a Intel(R) Xenon(R) CPU E5-2630 with 3.3 GHz clock speed (768 GB RAM) and 32 kernels. The linear and integer programming problems within CGGA are solved using the commercial solver Gurobi, version 8.1, restricted to maximal 6 kernels. We limit the computational time of each step in 2PH. Phase 1 is limited to 2 and 3 h for test instances 1 and 2, respectively. Phase 2 is limited to 9 h, thereof 3 h for the column generation procedure and 6 h for computing the integer solution.

We benchmark the test runs against a run without decomposition (CGGA) with the same time limits. For instance, the time limits of the column generation procedure and the computation of the integer solution for instance 1 are set to 5 h (equals the time limits of phase 1 plus the column generation procedure in phase 2) and 6 h, respectively.

Table 2 shows the minimum (*Min.*) and maximum (*Max.*) as well as the median (*Median*) of the best objective value found (*Obj.*) and its gap to the current lower bound (*Gap*) of 5 test runs per solution approach. We also include the share of mixed duties covering trips of more than one network (*MixD*) and the relative improvement of the objective value (Δ_{Obj}) in comparison to the non-decomposition benchmark. If the solution of the best objective value found includes infeasible duties, the value is marked with (*) and additional penalty cost (500,000 per infeasible duty) are added to the objective value. Test instance 1 consists of 2 networks. Hence, the number of sub-problems for all decomposition approaches is set to 2. As H1 and H2 produce the same decomposition, we only show the results for H1 in this case. Test instance 2 is decomposed into 4 sub-problems. Other experiments have shown that this number of sub-problems provides a good trade-off between solution time versus computational power consumption and solution quality in phase 1. The non-decomposition benchmarks could not produce feasible solutions in the given time limits. The best objective values found (with gaps) are $27,315,360^{(*)}$ (10.0%) and $10,171,075^{(*)}$ (4.1%) for test instance 1 and 2, respectively.

Table 2. Computational results of three decomposition strategies with test instances 1 and 2

	2PH_GP				2PH_H1				2PH_H2			
	Obj	Gap	MixD	Δ_{Obj}	Obj	Gap	MixD	Δ_{Obj}	Obj	Gap	MixD	Δ_{Obj}
Test instance 1												
Min	$26,237,050^{(*)}$	9.3%	52.5%	−3.9%	25,301,380	8.1%	30.9%	−7.37%	cf. 2PH_H1			
Median	$26,371,780^{(*)}$	9.6%	52.9%	−3.5%	25,614,250	9.3%	29.5%	−6.23%	cf. 2PH_H1			
Max	$27,024,370^{(*)}$	11.4%	54.5%	−1.1%	25,796,835	10.2%	29.7%	−5.56%	cf. 2PH_H1			
Test instance 2												
Min	$8,399,465^{(*)}$	5.2%	58.5%	−17.4%	7,728,480	4.4%	28.9%	−24.0%	7,679,230	5.0%	43.7%	−24.5%
Median	$8,427,005^{(*)}$	5.1%	62.3%	−17.2%	7,780,030	4.7%	30.1%	−23.5%	7,768,215	5.6%	52.4%	−23.6%
Max	$8,469,950^{(*)}$	5.4%	62.4%	−16.7%	$8,233,580^{(*)}$	4.2%	30.5%	−19.1%	7,809,230	6.0%	46.8%	−23.2%

It can be seen that all decomposition strategies lead to an improvement of the non-decomposition benchmark. However, similar to CGGA, the 2PH_GP approach does not achieve a feasible solution within the given time limits for both instances. In contrast, only one run of the network-based decomposition heuristics, the maximum of 2PH_H1 for test instance 2, produces an infeasible duty. All other runs lead to feasible solutions which indicates that maintaining the network structure during the decomposition supports the solution of the

MNRCSPAR. Moreover, the minimum, maximum and median of 2PH_H2 are slightly better than the comparable results of 2PH_H1 for instance 2 in terms of best objective found the related gaps. It also results in a higher share of mixed duties (43 up to 53%). Hence, considering the degree of connectivity between the networks further improves the already good results of the network-based decomposition without connectivity. Finally, the gaps of instance 1 of more than 8% indicate that great improvement of the best integer solution found is possible with longer computational times.

7 Conclusion

This paper presents a multi-network railway crew scheduling problem with application to the German regional passenger transportation and its specific requirement of attendance rates for conductors. We especially focus on the development of a problem-specific solution approach with the aim to achieve high quality solutions in reasonable time. Therefore, we develop a 2-phase optimization method which builds on a hybrid column generation and genetic algorithm, an existing solution approach for solving the single network case (see [8])). In the first phase, sub-problems are built and solved in parallel. By a variable selection mechanism, the duties of the solutions of the relaxed master problems are passed to the second phase in which the original problem is solved and an integer solution is generated.

We experiment with three different decomposition strategies: a graph partitioning algorithm developed by [14] with problem-specific adjusted edge weights and two variants of a network-based greedy decomposition heuristics. The decision rules build on the number of trips per network (variant 1) plus the degree of connectivity between networks (variant 2). We test the decomposition strategies with the 2-phase optimization method. The test results of two real-world instances show that maintaining the network structure by using the network-based greedy decomposition heuristics performs best within the given time limits. Hereby, considering the connectivity degree between the networks, next to keeping the number of trips across sub-problems balanced, achieves slightly better results than the decomposition without connectivity. Moreover, we show that feasible solutions for test instances with more than 18,000 trips can be generated within less than 12 h, but with potential for further improvement.

A disadvantage of 2PH_H1 and 2PH_H2 is that they are very problem-specific to the multi-network case and cannot be applied to a very large and complex single network. Still, they work well for our purpose. Further research involves testing the performance of 2PH_H2 with even larger instances, in particular planning test instance 2 with 11 networks for one week which is the standard planning period of the railway company. In this context, further solution method extensions with focus on accelerating the integer solution generation (e.g., variable fixing heuristics) should be investigated.

References

1. Abbink, E.J.W., Wout, J.V., Huisman, D.: Solving large scale crew scheduling problems by using iterative partitioning. In: Liebchen, C., Ahuja, R.K., Mesa, J.A. (ed.) ATMOS 2007. OASIcs 7, pp. 96–106. Dagstuhl Publishing, Saarbrücken (2007)
2. Caprara, A., Kroon, L., Monaci, M., Peeters, M., Toth, P.: Passenger railway optimization. In: Barnhart, C., Laporte, G. (eds.) Transportation, pp. 129–187. Elsevier, Amsterdam (2007)
3. Desaulniers, G., Desrosiers, J., Solomon, M.: Accelerating strategies in column generation methods for vehicle routing and crew scheduling problems. In: Ribeiro, C.C., Hansen, P. (eds.) Essays and Surveys in Metaheuristics, pp. 309–324. Springer, London (2002)
4. Fores, S., Proll, L., Wren, A.: Experiences with a flexible driver scheduler. In: Voß, S., Daduna, J.R. (eds.) Computer-Aided Scheduling of Public Transport, pp. 137–152. Springer, Heidelberg (2001)
5. Freling, R., Lentink, R.M., Odijk, M.A.: Scheduling train crews: a case study for the Dutch railways. In: Voß, S., Daduna, J.R. (eds.) Computer-Aided Scheduling of Public Transport, pp. 153–165. Springer, Heidelberg (2001)
6. Heil, J.: A solution approach for crew scheduling with attendance rates for multiple networks. In: Operations Research Proceedings. Springer (2018)
7. Huisman, D., Kroon, L.G., Lentink, R.M., Vromans, M.J.C.M.: Operations research in passenger railway transportation. Stat. Neerl. **59**(4), 467–497 (2005)
8. Hoffmann, K., Buscher, U., Neufeld, J.S., Tamke, F.: Solving practical railway crew scheduling problems with attendance rates. Bus. Inf. Syst. Eng. **59**, 147–159 (2017)
9. Hoffmann, K., Buscher, U.: Valid inequalities for the arc flow formulation of the railway crew scheduling problem with attendance rates. Comput. Ind. Eng. **127**, 1143–1152 (2019)
10. Ibarra-Rojas, O.J., Delgado, F., Giesen, R., Muñoz, J.C.: Planning, operation, and control of bus transport systems: a literature review. Transp. Res. Part B Methodol. **77**, 38–75 (2015)
11. Janacek, J., Kohani, M., Koniorczyk, M., Marton, P.: Optimization of periodic crew schedules with application of column generation method. Transp. Res. Part C Emerg. Technol. **83**, 165–178 (2017)
12. Jütte, S., Thonemann, U.W.: Divide-and-price: a decomposition algorithm for solving large railway crew scheduling problems. Eur. J. Oper. Res. **219**(2), 214–223 (2012)
13. Jütte, S., Thonemann, U.W.: A graph partitioning strategy for solving large-scale crew scheduling problems. OR Spectrum **37**(1), 137–170 (2015)
14. Karypis, G., Kumar, V.: A fast and high quality multilevel scheme for partitioning irregular graphs. SIAM J. Sci. Comput. **20**(1), 359–392 (1999)
15. Kasirzadeh, A., Saddoune, M., Soumis, F.: Airline crew scheduling: models, algorithms, and data sets. EURO J. Transp. Logistics **6**(2), 111–137 (2017)
16. Koniorczyk, M., Talas, B., Gedeon, F.: Preconditioning in the backtracking duty generation of passenger rail crew scheduling: a case study. Commun. Sci. Lett. Univ. Zilina **17**(2), 23–29 (2015)
17. Kwan, R.S.K., Kwan, A.S.K.: Effective search space control for large and/or complex driver scheduling problems. Ann. Oper. Res. **155**(1), 417–435 (2007)
18. Lübbecke, M.E., Desrosiers, J.: Selected topics in column generation. Oper. Res. **53**(6), 1007–1023 (2005)

19. Vaidyanathan, B., Jha, K.C., Ahuja, R.K.: Multicommodity network flow approach to the railroad crew-scheduling problem. IBM J. Res. Dev. **51**, 325–344 (2007)
20. Wang, Y., Liu, J., Miao, J., Nie, L.: A column generation-based approach for railway crew scheduling problem. In: Tian, Z., Huang, H., Gao, Z. (eds.) Mao B, pp. 697–708. Traffic and Transportation Studies, American Society of Civil Engineers (2008)

An Improved LP-Based Heuristic for Solving a Real-World Locomotive Assignment Problem

Martin Scheffler, Michael Hölscher, and Janis S. Neufeld(✉)

Fakultät Wirtschaftswissenschaften, Lehrstuhl für BWL, insb. Industrielles Management, Technische Universität Dresden, 01062 Dresden, Germany
{martin.scheffler,janis_sebastian.neufeld}@tu-dresden.de

Abstract. The locomotive assignment (or scheduling) problem is a highly relevant problem in rail freight transport. For a preplanned train schedule, minimum-cost locomotive schedules have to be created so that each train is pulled by the required number of locomotives (locomotives are assigned to trains). Determining locomotive schedules goes hand in hand with determining the number of required locomotives and this has a significant impact on capital commitment costs. Therefore, this paper proposes an improved heuristic for scheduling locomotives at a European rail freight operator. We show that a transformation of an iterative process to simplify the underlying network into a one-step procedure can significantly reduce computing times of a heuristic. Computational tests are carried out on the real-world instance as well as on smaller instances. The results show that the proposed heuristic outperforms an existing heuristic from literature in terms of both solution quality and computation times and, in contrast to approaches from literature, enables a solution of a practical instance in Europe.

Keywords: Locomotive assignment · Linear programming · Heuristic

1 Introduction

The railway sector in Europe is characterized by a strong competition and a continuing privatization trend. It is therefore important for the railway companies to exploit cost saving potentials to keep their competitiveness [4]. Because of the high costs for the operation of trains and the high acquisition costs for locomotives an efficient deployment of rolling stock is strongly relevant [11]. This is mainly determined by the so-called Locomotive Assignment Problem (LAP), which describes the assignment of locomotives to trains (preplanned train schedule) with consideration of several side constraints [8]. Both operational costs (e.g. petrol/electricity) and fixed costs for the use of locomotives are minimized. Even small improvements with regard to an efficient locomotive assignment can lead to significant economic savings, while at the same time the LAP is a highly complex planning problem [6].

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, pp. 314–329, 2019.
https://doi.org/10.1007/978-3-030-29821-0_21

Various approaches to tackle the LAP have been presented in literature, which are summarized by [7]. Some studies focus on locomotive or rolling stock scheduling in passenger transport (see e.g. [10]). However, there are several characteristics that make the developed algorithms difficult to apply to freight transport, such as the timely distribution of trains, technical restriction or the different planning procedures in passenger and freight transport. In general, LAP approaches can be classified in single locomotive models and multiple locomotive models [7]. While in single locomotive models only one locomotive is assigned to each train (see e.g. [2,5]), in multiple locomotive models a combination of two or more locomotives, are formed and scheduled [3,9,12,13]. Such a combination of several locomotives is called consist. Often consists are necessary to gain enough engine power for pulling the designated trains. With this, the resulting LAP is much more difficult to solve. In addition, several types of locomotives may exist in practical problems, which is regarded by [1], who study a LAP in North America. This increases complexity as well, but the consideration is often necessary to provide real-world decision support. For the practical LAP studied in this paper, only these approaches are suitable. Still, it has to be noted that the LAP in European freight transport differs significantly from problems in North America. In particular, in Europe relatively short trains are prevalent, while, at the same time, the number of trains is rather large. Moreover, as a train may have to cross several borders, not all locomotives might meet legal or technical conditions for pulling a train from end to end.

In this paper, we apply the heuristic by [1] to a real-world LAP for railway freight transport in Europe. Our results reveal that the existing approach is not able to solve large real-world problems with the characteristics mentioned above. Therefore, we propose an improved algorithm, which is based on a MIP formulation as multicommodity flow problem and, most notably, speeds up a heuristic step of simplifying the underlying network. Two variants are tested and compared to the existing approach.

Our work is structured as follows. In Sect. 2.1 we describe the specific characteristics of the studied LAP. The MIP formulation presented in Sect. 2.2 forms the basis for the presented heuristic in Sect. 3. Computational results for several test instances are discussed in Sect. 4, which is followed by a summary and outlook on fruitful directions for future research.

2 Problem Description

2.1 Characteristics of the Studied LAP

The locomotive assignment problem is modeled as a multicommodity flow problem with consideration of several side constraints as in [1]. In general the main objective is to find a feasible flow for certain commodities in a network while minimizing the total costs. In the context of the LAP, the commodities are represented by different locomotives. The task is to find an optimized assignment to trains and plan the flow of locomotives through the underlying network. The train schedule itself is planned in a preceding step and is therefore a direct input to the LAP. The planning horizon is given by one week.

The used network is modeled as a space-time network that represents the basis for the following optimization processes. A graphical illustration is shown in Fig. 1.

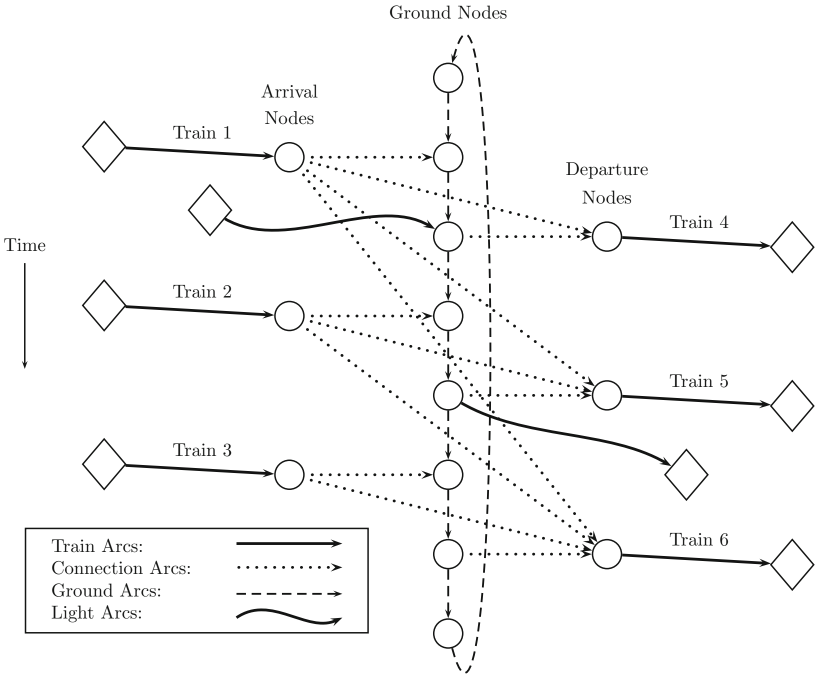


Fig. 1. Space-time network

The network consists of a set of nodes N that can be divided in ground nodes (N^{Ground}), departure nodes ($N^{Departure}$) and arrival nodes ($N^{Arrival}$). Nodes with a round shape are representing the same station. Angular nodes represent any other stations. The nodes are linked by a set of arcs A, that contains train arcs (A^{Train}), connection arcs ($A^{Connect}$), ground arcs (A^{Ground}) and light arcs (A^{Light}). The nodes and arcs are characterized by attributes for place and time. An important component of the network are the train arcs, which map the actual trains. These arcs connect a departure node and an arrival node of the respective train. For every event (departure and arrival of a train) ground nodes are added to the network with a relation to a specific train station. The connections between arrival or departure nodes and the corresponding ground nodes are implemented by connection arcs. A special subset of the connection arcs are the train-train connections ($A^{TrainTrain}$). These represent the direct linking between an arrival node and a departure node of a later train that leaves from the same train station. In this case, the whole consist, that is pulling an incoming train, remains unaffected and is transferred without any changes in its structure to a later outgoing train. An alternative is offered by so called consist busting, that describes the process of splitting up a consist of locomotives and regrouping them for other trains. In this case, the locomotives use an connection arc from the arrival node to the corresponding ground node (A^{Bust}) and the locomotives get to a pool of vehicles that could be composed to new consists for upcoming trains. The ground nodes are connected by ground arcs for modelling idle times of locomotives at the respective train stations, whereas the last ground node (N^{Last}) of the planning horizon at each station has an outgoing ground arc to the first ground node to create a cyclic assignment plan. If this is ensured, the closing balance of the current time period matches with the opening balance of the next time period in terms of the number of locomotives at each station.

The locomotives are given by set L. Since each type of locomotive l can only drive on a part of the rail network (e.g. electric locomotives cannot be assigned to a train that goes along a non-electrified track), a set A_l is introduced, which contains all passable arcs for locomotive type l. For the locomotives in the space-time network, there are three different possibilities of moving. First, locomotives are able to actively pull trains on a train arc. Moreover, locomotives could attend other locomotives passively, that means they are pulled by other locomotives. This process is called deadheading. Finally, locomotives can use light-traveling to move to other train stations. In this case, the locomotives do not pull a train but independently change their location in the network. Light-traveling can be used, for example, to balance availability of locomotives at train stations. Basically, a light arc (A^{Light}) always connects two ground nodes of different stations. Like in [1] we create light arcs departing with the fixed time interval of eight hours at a train station. This means it is possible to reach every other station by light-traveling every eight hours. However, we illustrate only two examplary light arcs in Fig. 1 to ensure clarity.

2.2 Mathematical Formulation

We present a linear mixed integer programming model for the multicommodity flow problem based on [1]. Note that constraints (2)–(7) are directly adapted. The objective function and constraints (8)–(18) are different. Table 1 displays the used notation.

The objective function (1) of the MIP formulation describes the total cost function of the LAP, that includes several terms representing the influencing factors. These are, firstly, the costs of active locomotives pulling trains on train arcs (γ_{la}^{Active}) associated with the number of locomotives flowing on these arcs (x_{la}). The second term describes the costs of deadheading locomotives ($\gamma_{la}^{Passive}$) multiplied with the amount of these locomotives (y_{la}) flowing on the corresponding train arcs. The costs for light-traveling are modeled analogously and are calculated by the product of the cost rate (γ_{la}^{Light}) and the number of flowing locomotives on the light arcs (y_{la}). Note that for active pulling and light travelling set A_l (passable arcs) is taken into account. Note y_{la} is used for modelling

Table 1. Notation in MIP-Formulation

Sets	
A	Set of all arcs a
A^{Bust}	Subset of $A^{Connect}$ with all arcs a causing consist busting
$A^{Connect}$	Set of all arcs a representing connection arcs
A^{Ground}	Set of all arcs a representing ground arcs
A_i^{In}	Set of all incoming arcs a at node i
A_l	Subset of A with all passable arcs a for locomotive type l
A^{Last}	Subset of A^{Ground} with all arcs a ending in the temporally last ground node of a time period at each station
A^{Light}	Set of all arcs a representing light-traveling
A_i^{Out}	Set of all outgoing arcs a at node i
A^{Train}	Set of all arcs a representing a train
$A^{TrainTrain}$	Subset of $A^{Connect}$ with all arcs representing train-train connections
L	Set of all locomotive types l
N	Set of all nodes i
$N^{Arrival}$	Set of all nodes i with incoming train arcs
$N^{Departure}$	Set of all nodes i with outgoing train arcs
Parameters	
γ_{la}^{Active}	Cost of an active locomotive of type l on train arc a
γ^{Bust}	Cost of consist busting
γ_l^{Fix}	Cost of using one locomotive of type l
γ_{la}^{Light}	Cost of light-traveling on light arc a
$\gamma_{la}^{Passive}$	Cost of deadheading locomotive of type l on arc a
γ^{Pen}	Penalty costs
B_l	Number of available locomotives of type l
K	Maximum number of locomotives on an arc
T_a	Tonnage requirement on train arc a
t_{la}	Tonnage pulling capability of locomotive type l on train arc a
Variables	
s_l	Integer variable, number of locomotives of type l exceeding B_l
u_l	Integer variable, number of used locomotives of type l
v_a	Binary variable, 1 if at least one locomotive flows on arc a
x_{la}	Integer variable, number of active locomotives of type l on train arc a
y_{la}	Integer variable, number of locomotives not pulling a train (including deadheading, light-traveling, idling) of type l on arc a

different things on different arcs: deadheading on train arcs, light-travelling on light arcs and the general flow on ground and connection arcs. The fourth term of the objective function describes consist busting. Therefore the corresponding costs (γ^{Bust}) are multiplied with the binary decision variable v_a, that becomes 1 if at least one locomotive flows on arc a of the set A^{Bust}. The last two terms of the objective function describe the fixed costs for using locomotives of different types (γ_l^{Fix}) multiplied with the number of used locomotives (u_l) and the penalty costs for exceeding the available number of locomotives (γ^{Pen}) by the amount of s_l locomotives. This proceeding differs from the proposed one of [1], who work with a formulation that only includes the possibility of saving a certain number of locomotives in comparison with the fleet size B_l. But to preserve the feasibility of the model for fictive or unknown instances, where the actual fleet size B_l can be uncertain or needs to be set manually, this paper considers a formulation where the used locomotives u_l are counted and B_l can be exceeded by the integer variable s_l. In practice, this feature makes it possible to create a solution at any time. This enables the planner to analyze the reason for the inadmissibility.

Constraints (2) ensure that the assigned locomotives are able to pull the weight of the trains. The following constraints (3) of the MIP formulation ensure that not more than K locomotives are used on train arcs and light arcs. Constraints (4) control the flow balance and state that at each node the number of incoming locomotives has to be equal to the number of outgoing locomotives. Constraints (5) assign the value 1 to the variable v_a if a locomotive flows on the connection arc a, which is needed for the calculation of fixed costs for consist busting. Constraints (6) ensure that at each arrival node all the involved locomotives use only one outgoing connection arc either to a ground node (causing consist busting) or to a subsequent departure node (train-train connection). Similarly constraints (7) state that all locomotives leaving a departure node are flowing on the same incoming connection arc before. With constraints (8) the number of used locomotives is stored in the variable u_l and constraints (9) maintain the described opportunity of exceeding the fleet size B_l by using the penalized variable s_l to guarantee a feasible solution. The constraints (10)–(18) describe the type of the different decision variables and their definition and value range.

$$\begin{aligned}\min z = &\sum_{l\in L}\sum_{a\in A^{Train}\cap A_l}\gamma_{la}^{Active}\cdot x_{la} + \sum_{l\in L}\sum_{a\in A^{Train}}\gamma_{la}^{Passive}\cdot y_{la}\\ &+\sum_{l\in L}\sum_{a\in A^{Light}\cap A_l}\gamma_{la}^{Light}\cdot y_{la} + \sum_{a\in A^{Bust}}\gamma^{Bust}\cdot v_a\\ &+\sum_{l\in L}\gamma_l^{Fix}\cdot u_l + \sum_{l\in L}\gamma^{Pen}\cdot s_l \end{aligned} \tag{1}$$

$$\sum_{l \in L: a \in A_l} t_{la} \cdot x_{la} \geq T_a \qquad \forall a \in A^{Train}, \tag{2}$$

$$\sum_{l \in LL: a \in A_l} (x_{la} + y_{la}) \leq K \forall a \in A^{Train}, \tag{3}$$

$$\sum_{a \in A_i^{In} \cap A_l} x_{la} + y_{la} = \sum_{a \in A_i^{Out} \cap A_l} x_{la} + y_{la} \qquad \forall i \in N, \forall l \in L, \tag{4}$$

$$\sum_{l \in L} y_{la} \leq K \cdot v_a \qquad \forall a \in A^{Connect}, \tag{5}$$

$$\sum_{a \in A_i^{Out}} v_a = 1 \qquad \forall i \in N^{Arrival}, \tag{6}$$

$$\sum_{a \in A_i^{In}} v_a = 1 \qquad \forall i \in N^{Departure}, \tag{7}$$

$$u_l = \sum_{a \in A^{Last}} y_{la} \qquad \forall l \in L, \tag{8}$$

$$u_l \leq B_l + s_l \qquad \forall l \in L, \tag{9}$$

$$s_l, u_l \in \mathbb{N} \qquad \forall l \in L, \tag{10}$$

$$s_l, u_l \geq 0 \qquad \forall l \in L, \tag{11}$$

$$v_a \in \{0, 1\} \qquad \forall a \in A^{Connect}, \tag{12}$$

$$x_{la} \geq 0 \qquad \forall l \in L, \forall a \in A^{Train} \cap A_l, \tag{13}$$

$$y_{la} \geq 0 \qquad \forall l \in L, \forall a \in A^{Light} \cap A_l, \tag{14}$$

$$y_{la} \geq 0 \qquad \forall l \in L, \forall a \in A \setminus A^{Light}, \tag{15}$$

$$x_{la} \in \mathbb{N} \qquad \forall l \in L, \forall a \in A^{Train} \cap A_l, \tag{16}$$

$$y_{la} \in \mathbb{N} \qquad \forall l \in L, \forall a \in A^{Light} \cap A_l, \tag{17}$$

$$y_{la} \in \mathbb{N} \qquad \forall l \in L, \forall a \in A \setminus A^{Light}. \tag{18}$$

3 Solution Approaches

3.1 Complexity of the Mathematical Model

While analyzing the problem structure of the LAP, it can be observed that the number of arcs is increasing rapidly for larger instances. The size of the solution space is mainly affected by the relation between the number of train stations, the number of scheduled trains and the number of different locomotive types, that could be used for the assignment. [1] have shown that the LAP is NP-complete. Therefore, it is necessary to develop heuristic methods to solve the LAP or to reduce the associated solution space. An important starting-point to reducing the complexity of the problem are the train-train connection arcs (subset of $A^{Connect}$), that link arrival nodes with departure nodes in the space-time network. The amount of connection arcs (specifically train-train connections) usually exceeds the number of train arcs by far and is strongly dependent of the specific problem instance and the scheduled trains. In addition, the proposed MIP formulation in combination with all the possible train-train connections leads to a large number of binary variables (v_a), that cause an increasing problem complexity. Because of constraints (6) and (7) of the model, one needs to solve a decision problem at every arrival (and departure) node, i.e. which of the outgoing (incoming) arcs should be used, as on only one of them has to be a positive flow of locomotives. This is necessary because the used consist is either busted (consist busting arc is used) and the locomotives flow to a ground node or they are transferred altogether to another departure node (train-train connection arc is used). It is not possible to split the consist and send the locomotives on different train-train connections or have a mixture of consist busting and the usage of train-train connections at a specific arrival (departure) node. Otherwise, the consist busting costs could not be accurately mapped.

To reduce the amount of connection arcs in the space-time network and to accelerate the solution process it is reasonable to shorten the time window for feasible train-train connections in a first step [1]. We have defined that it is only possible to use a train-train connection and transfer a consist of locomotives to an other departure node in the first 12 h after the arrival at a train station. This avoids very long idle times, as it can be assumed that they do not contribute to a minimum-cost schedule.

3.2 LP-Based Heuristics

For further reduction of arcs in the space-time network, [1] propose an iterative heuristic method to determine certain train-train connections. Although our formulation is very similar, but not identical with [1], the heuristic can still be adapted without changes. Firstly, the integrality restrictions are removed to obtain the linear programming relaxation of the original MIP model. Furthermore, the binary variable v_a and the associated constraints (5)–(7) are eliminated from the model formulation. With this, it is possible that on more than one outgoing (incoming) arc at an arrival (departure) node locomotives are flowing, which reduces the solution time significantly. In addition, high costs are assigned to all arcs of the set A^{Bust} (train-to-ground connection arcs) to prevent consist busting and force the flowing locomotives on the relevant train-train connection arcs. For our problem the result is a new objective function shown by (1').

$$\min z = \sum_{l \in L} \sum_{a \in A^{Train} \cap A_l} \gamma_{la}^{Active} \cdot x_{la} + \sum_{l \in L} \sum_{a \in A^{Train} \cap A_l} \gamma_{la}^{Passive} \cdot y_{la} \\ + \sum_{a \in A^{Light} \cap A_l} \sum_{l \in L} \gamma_{la}^{Light} \cdot y_{la} + \sum_{a \in A^{Bust}} \sum_{l \in L} \gamma^{Pen} \cdot y_{la} \\ + \sum_{l \in L} \gamma_l^{Fix} \cdot u_l + \sum_{l \in L} \gamma^{Pen} \cdot s_l \qquad (1')$$

In summary, the linear programming relaxation is given by min (1') s.t. (2)–(4), (8)–(15). In a following step the cumulated flow of locomotives on potential train-train connection arcs $a \in A^{Candidate}$ is calculated and represented by the variable $\varphi(a)$. The train-train connection arc with the largest value of $\varphi(a)$ seems to be a good choice for a train-train connection and is determined to be the only possible connection between the two associated train arcs. Afterwards, the linear programming relaxation is solved again and the related objective value is analyzed. If it has increased by an amount that exceeds the parameter θ, the fixation of the train-train connection is reversed, otherwise it is kept. θ is used here as the treshold for worsening the objective value and is assumed to be 1000, analogous to [1]. This iterative procedure is repeated until either no more potential train-train connections ($A^{Candidate}$) are left or until a specified number of determined train-train connections represented by parameter γ has been reached. For the experimental tests (see Sect. 4) we have changed these termination criteria to a simple time based termination. The pseudo-code of the algorithm is given in Algorithm 1.

Algorithm 1.
Original heuristic presented by [1]

$A^{Candidate} = A^{TrainTrain}$; $A^R = \emptyset$
while time limit not reached & $A^{Candidate} \neq \emptyset$ **do**
  min (1') s.t. (2)-(4), (8)-(15).
  if current objective $\leq$ previous objective + θ **then**
    $A^{tmp} = \emptyset$
    determine $\varphi(a)$ $\forall a \in A^{Candidate}$
    choose a^* to which applies $\forall a \in A^{Candidate} : \varphi(a) \leq \varphi(a^*)$
    remove a^* from $A^{Candidate}$
    make a^* the only connection between the two associated train arcs
    add eliminated arcs to A^R and A^{tmp}
  else
    $A^R = A^R \setminus A^{tmp}$
for all $l \in L, a \in A^R$ **do**
  add y_{la} to Y^R
min (1) s.t. (2)-(18), $y_{la} = 0 \; \forall \, y_{la} \in Y^R$.

From our perspective, one main disadvantage of this heuristic is that especially for larger instances the linear programming relaxation has to be solved very often to determine a sufficient number of train-train connection arcs. Even for the relaxed model the solution time is increasing significantly with growing problem size. Hence, it is unfavorable to solve it very often, which is necessary for real-world instances. Moreover, it could happen that after one iteration the preceded fixation has to be reversed, which affects the efficiency of the heuristic. Therefore, we propose an improved heuristic to determine train-train connection arcs for the LAP, which is described in the following.

To avoid the determination of only one train-train connection per iteration, the linear programming relaxation is only solved once in our approach. After that, the flow of locomotives on train-train connections arcs $\varphi(a)$ is calculated and all of the arcs with $\varphi(a) = 0$ are removed from the set of connection arcs in a single step. In the first version of the heuristic, the cumulated flow of all types of locomotives (CF) is considered at that point, see Algorithm 2. We use $\bar{y}_{la}$ as solution of solving the linear programming relaxation and set Y^R to store removed variables temporarily. In contrast, in a second version (see Algorithm 3) the flow of each individual locomotive type (SF) is taken into account and the train-train connection arcs are blocked for specific types of locomotives. If only train-train connection arcs with a flow of zero locomotives are removed from the network, for both variants it is guaranteed that the objective function value of the relaxed model is not affected. That makes a change analysis of the objective function value like in the algorithm of [1] unnecessary. Since the objective value is not affected by removing arcs/variables with zero flow, we are able to remove all in one step. With this new procedure it is possible to reduce the number of arcs in the underlying space-time network quickly and preserve only promising candidates for train-train connections. To determine the same amount of train-train connections with the approach of [1] significantly more time is necessary. Furthermore, after using the heuristic of [1], consist busting is taking place more often, because of the small number of determined train-train connections. Hence, the cost-saving potential is probably not exploited in the same extent as in the proposed improved algorithm.

Algorithm 2.
Cum. flow based heuristic (CF)

min (1') s.t. (2)-(4), (8)-(15).
for all $a \in A^{TrainTrain}$ **do**
    $\varphi(a) = \sum_{l \in L} \bar{y}_{la}$
    if $\varphi(a) = 0$ **then**
        for all $l \in L$ **do**
            add y_{la} to Y^R
min (1) s.t. (2)-(18), $y_{la} = 0 \; \forall \; y_{la} \in Y^R$.

Algorithm 3.
Single flow based heuristic (SF)

min (1') s.t. (2)-(4), (8)-(15).
for all $a \in A^{TrainTrain}$ **do**
    for all $l \in L$ **do**
        if $\bar{y}_{la} = 0$ **then**
            add y_{la} to Y^R
min (1) s.t. (2)-(18), $y_{la} = 0 \; \forall \; y_{la} \in Y^R$.

4 Experimental Tests

4.1 Experimental Design

The heuristic of [1] as well as the improved heuristics described in Sect. 3 were implemented in `C#`. For solving the optimization problems we used Gurobi (8.0.0). All tests were run with a limit of 4 parallel threads on an Intel(R) Xeon(R) CPU E5-2630 v2 with 2.6 GHz clock speed and 384 GB RAM.

As mentioned in Sect. 1, the solution approaches were used to solve a real-world instance of an European rail freight operator. The instance represents a train schedule for one week which covers four central European countries. 2342 trains are distributed over 121 track sections. The sections connect 76 stations. The total number of locomotives is 162 with 13 different locomotive types. In order to be able to guarantee meaningful tests, we derived smaller test instances of different sizes from the real-world problem. This can easily be done by using network-specific knowledge to ignore single trains or groups of trains. The important set sizes for all instances are summarized in Table 2. The small (middle, large, very large) instances are denoted by 's' ('m', 'l', 'v') and are numbered consecutively. Instance 'r' is the real-life instance. Note that the set of light arcs A^{Light} is the same size for all instances (41012) because we use the same creating procedure like [1]. Light arcs are created every 8 h between all stations.

Table 2. Set sizes of the considered instances

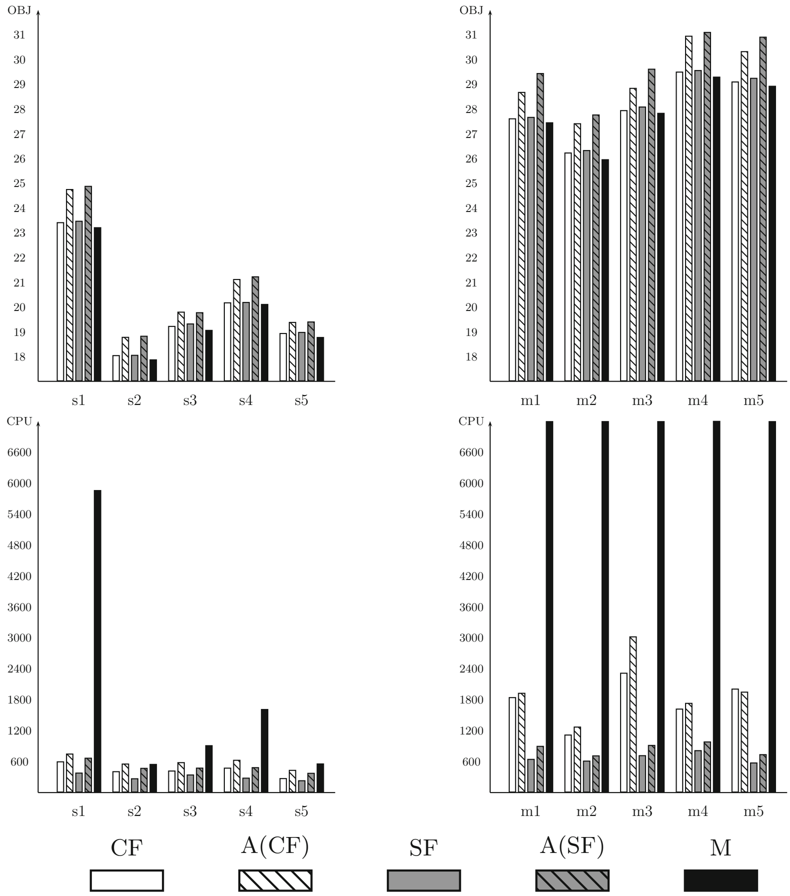
	N	A	A^{Train}	A^{Train}_{Train}
s1	44653	91115	831	4646
s2	43815	87870	612	2458
s3	43888	88769	632	3264
s4	44096	89205	681	3443
s5	43759	87895	594	2557
l1	47662	107021	1615	16759
l2	46988	102479	1444	13062
l3	47129	104262	1477	14671
l4	47322	104788	1526	14955
l5	47103	102654	1468	13098
r	50422	128201	2342	34452

	N	A	A^{Train}	A^{Train}_{Train}
m1	45646	95300	1090	7579
m2	45366	94689	1016	7322
m3	45601	95519	1072	7861
m4	45709	95805	1104	8007
m5	45906	96245	1158	8196
v1	49105	117746	2000	25656
v2	48780	114236	1913	22558
v3	48987	115143	1962	23209
v4	49063	117606	1982	25576
v5	48231	111238	1763	20259

To keep the number of tests manageable, in a first step we compare both variants of the improved heuristic (CF - cumulated flow based; SF - single flow based) and the heuristic of [1] (A) with solution of the original MIP formulation presented in Sect. 2.2 (M). For this, we use only the small and medium-sized instances and limit the computing time to two hours. Afterwards, the best variant of our heuristic (CF or SF) is compared to [1] for the large and very large instances. Finally, we solve the real-life instance with a time limit of six hours.

4.2 Experimental Results

Figure 2 shows the results for comparing the improved heuristic (CF and SF) with the heuristic of [1] (A) and the original MIP formulation (M). The presented values are averages of 5 runs. Each MIP was terminated with a gap less or equal 1%. In order to allow a fair comparison, A was solved twice (A(CF) and A(SF))

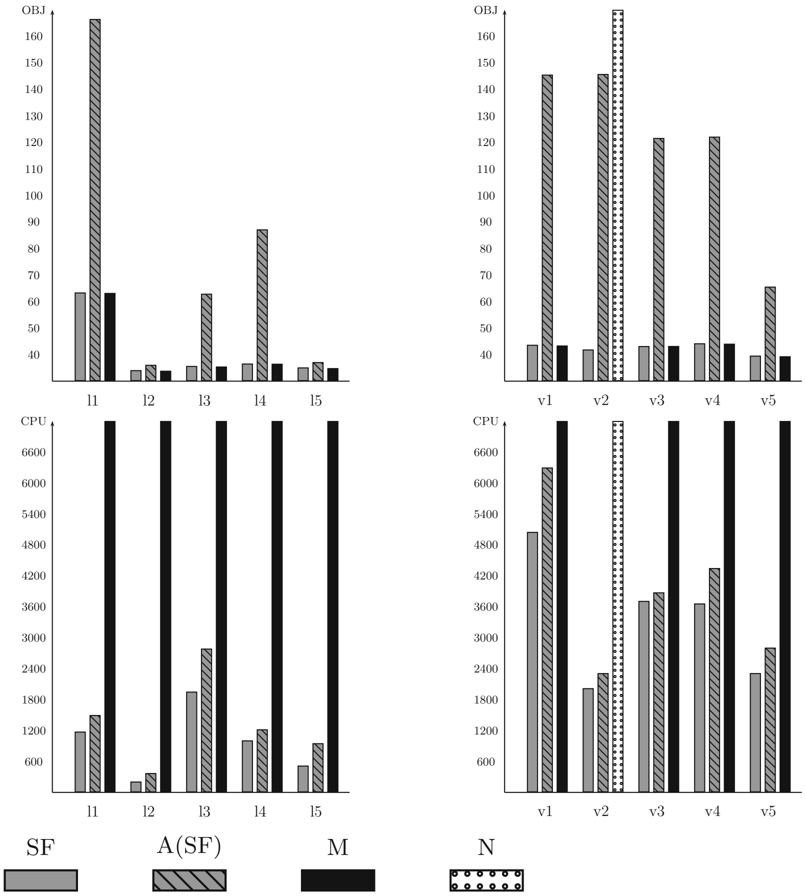


Notation: CF: cumulated flow based heuristic; SF: single flow based heuristic; A(CF): heuristic of [1] with computing time limited to maximum of CF; A(SF): heuristic of [1] with computing time limited to maximum of SF; M: original MIP formulation; OBJ in millions; CPU in seconds

Fig. 2. Results for the small and mid sized instances

with different time limits. The time limits were set based on the slowest runs of the corresponding improved heuristics (CF or SF). Furthermore, the available time for A has to be split into two parts: iterative procedure and solving the reduced MIP. Preliminary tests show that using 75% of time for the former and 25% for the latter is suitable.

It can be seen that the improved heuristics achieve significantly better objective values than A. CF achieves slightly better values than SF, since the solution space is less restricted by the heuristic. However, this small disadvantage is compensated by the considerably faster computing time. This in turn, is caused

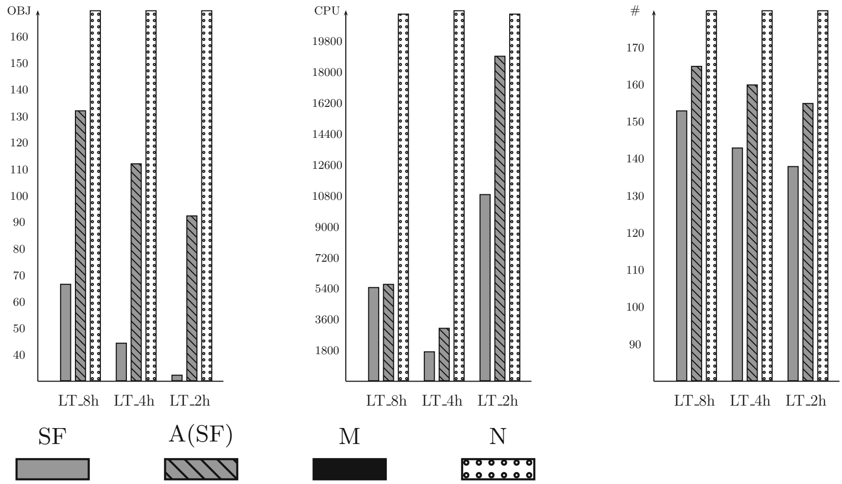


Notation: SF: single flow based heuristic; A(SF): heuristic of [1] with computing time limited to maximum of SF; M: original MIP formulation; N: no solution gained by M; OBJ in millions; CPU in seconds

Fig. 3. Results for the large and very large sized instances

by the smaller remaining solution space. Thus, we decided to use only SF as improved heuristic for all further tests. For these instance sizes (nearly) optimal solutions can be obtained in reasonable time by solving the original MIP formulation (M). The convergence speed on the medium-sized instances is sufficient to keep up with CF and SF. On average, a gap of less than 5% (2%) could be achieved in less than 5 (45) min.

Figure 3 shows the results for the large and very large instances. Again, solving the original MIP formulation leads on average to solutions with a gap of less than or equal to 5% within about half an hour. However, for instance v2 not every run of M could create a feasible solution at all (marked as N in Fig. 3). Once more SF produces much better results than A within the same time. Moreover, SF is



Notation: SF: single flow based heuristic; A(SF): heuristic of [1] with computing time limited to maximum of SF; M: original MIP formulation; N: no solution gained by M; OBJ in millions; CPU in seconds, # number of locomotives

Fig. 4. Results for the real-life instances

able to create average optimality gaps of about 3%. These gaps are calculated in relation to the highest lower bound obtained by all runs of M. In addition, A is not able to generate practically executable solutions for all instances. This means for instances l1, l3, l4 and all very large instances more locomotives are scheduled then available. However, this also applies to SF and M on instance l1.

The reason for this is the used time interval for creating light arcs. As mentioned in Sect. 2.1, we have chosen the value of eight hours analogously to [1]. However, the set of light arcs is not sufficient to create admissible schedules. Therefore, it is necessary to vary this parameter for the last test. Figure 4 shows the results for the real-life instance with a variation of light-traveling every eight, four and two hours (LT_8h, LT_4h, LT_2h). In addition, we have also shown the number of locomotives required in the final schedule. Note that even though solutions for LT_4 and LT_8 do not exceed the total number of existing locomotives, this is still the case for individual locomotive types. It can be seen that M is not able to generate valid solutions for all frequencies of light-traveling. Again, SF outperforms A for each setting. Only SF is able to generate admissible solutions with light arcs every two hours. Optimality gaps can be obtained by using the highest lower bound generated by all runs of M. The average gap of SF is lower than 7%. In addition, SF enables us to solve this real-life instance of an European rail freight operator within reasonable time for the tactical planning level. A computing time of about 1 h (LT_2h) can be assessed positively.

5 Conclusions

In this paper, we studied a practical LAP of a European railway freight transport company. Since existing approaches were not able to solve the real-world instance, we proposed an improved heuristic, that determines train-train connections efficiently and is therewith able to simplify the underlying solution space effectively. By comparing to variants of the algorithm for several test instances a preferable algorithm could be identified, that is also able to solve the real-world instance within reasonable time.

The small number of approaches that are suitable to solve complex practical locomotive assignment problems and integrate necessary restrictions indicates a big potential for future research. Some of these restrictions are, for example, the fact that not all locomotive types can be combined with each other or a detailed modelling of unavoidable times that occur through (dis-)connecting processes between cars and locomotives. A crucial point for future work is the removal of all used heuristic limitations of the solution space (e.g. light arcs are only available every 8 h, train-train connections limited to 12 h). As the results for the real instance have shown, it is important to find an automated and general approach for light-travelling. Furthermore, an integrated approach considering all factors that increase complexity is necessary. Nevertheless, we could show that with the improved heuristic real problem sizes can already be solved with high solution quality.

References

1. Ahuja, R.K., Liu, J., Orlin, J.B., et al.: Solving real-life locomotive-scheduling problems. Transp. Sci. **39**(4), 503–517 (2005)
2. Cordeau, J.F., Desaulniers, G., Lingaya, N., Soumis, F., Desrosiers, J.: Simultaneous locomotive and car assignment at VIA Rail Canada. Transp. Res. Part B Methodol. **35**, 767–787 (2001)
3. Florian, M., Bushell, G., Ferland, J., Guerin, G., Nastansky, L.: The engine scheduling problem in a railway network. INFOR Inf. Syst. Oper. Res. **14**(2), 121–138 (1976)
4. Hoffmann, K., Buscher, U., Neufeld, J.S., Tamke, F.: Solving practical railway crew scheduling problems with attendance rates. Bus. Inf. Syst. Eng. **59**(3), 147–159 (2017)
5. Lübbecke, M.E., Zimmermann, U.T.: Engine routing and scheduling at industrial in-plant railroads. Transp. Sci. **37**(2), 183–197 (2003)
6. Piu, F.: A mixed integer programming approach to the locomotive assignment problem. In: Swiss Transport Research Conference-(STRC), pp. 1–42 (2011)
7. Piu, F., Speranza, M.G.: The locomotive assignment problem: a survey on optimization models. Int. Trans. Oper. Res. **21**(3), 327–352 (2014)
8. Piu, F., Kumar, V.P., Bierlaire, M., et al.: Introducing a preliminary consists selection in the locomotive assignment problem. Transp. Res. Part E Logist. Transp. Rev. **82**, 217–237 (2015)
9. Powell, W.B., Topaloglu, H.: Stochastic programming in transportation and logistics. In: Handbooks in Operations Research and Management Science, vol. 10, pp. 555–636 (2003)

10. Reuther, M., Schlechte, T.: Optimization of rolling stock rotations. In: Handbook of Optimization in the Railway Industry, pp. 213–241 (2018)
11. Rouillon, S., Desaulniers, G., Soumis, F.: An extended branch-and-bound method for locomotive assignment. Transp. Res. Part B Methodol. **40**(5), 404–423 (2006)
12. Vaidyanathan, B., Ahuja, R.K., Orlin, J.B.: The locomotive routing problem. Transp. Sci. **42**(4), 492–507 (2008)
13. Ziarati, K., Soumis, F., Desrosiers, J., Gelinas, S., Saintonge, A.: Locomotive assignment with heterogeneous consists at CN North America. Eur. J. Oper. Res. **97**, 281–292 (1997)

Author Index

C. Bierwirth et al. (Eds.): *Logistics Management*, LNLO, p. 331, 2019.
https://doi.org/10.1007/978-3-030-29821-0