Logistics Research (2022) 15:7
DOI_10.23773/2022_7

Supply Chain Analytics:
Investigating Literature-Practice Perspectives and Research Opportunities

S. Lodemann[1], S. Lechtenberg[2], K. Wesendrup[2], B. Hellingrath[2], K. Hoberg[3], W. Kersten[1]

Received: 30 June 2021 / Accepted: 15 June 2022 / Published online: 20 July 2022

ABSTRACT

Supported by ever-increasing amounts of data and maturing technologies, big data analytics offers viable, promising improvements for various fields and applications. Supply chain analytics (SCA), the application of big data analytics to supply chain management, can enhance and innovate supply chain processes and services in most companies. To reap such benefits, supply chain managers must overcome various obstacles, including the identification of appropriate methods, data, and application cases. The degree to which the potential value of SCA actually is being harnessed by practitioners remains uncertain. The study aims to synthesize scientific and practical perspectives regarding the SCA dimensions: goal and motivation, method, data, and application area. For this purpose the research applies a multi-vocal literature review (MLR) and a survey approach. The study reviews over 1481 publications and consults 278 respondents to reveal six different goals and seven motivations for SCA. Moreover, descriptive, predictive, and prescriptive analytics and many different data types enabling SCA within different application areas are examined. The cross-analysis between scientific and practical perspectives identifies several gaps, such as lack of specific data usage, low practical SCA maturity, or undersaturated research areas that show future paths of academic research.

KEYWORDS: Supply Chain Analytics · Data Analytics · Supply Chain Management · Systematic Literature Review · Multivocal Literature Review

'''This article is part of a focus collection on ''Supply Chain Analytics in the 2020s".

Sebastian Lodemann[1], Corresponding author
sebastian.lodemann@tuhh.de, +49 40 42878 4505

Sandra Lechtenberg[2]

Kevin Wesendrup[2]

Bernd Hellingrath[2]

Kai Hoberg[3]

Wolfgang Kersten[1]

[1] Hamburg University of Technology;
Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany

[2] University of Münster;
Leonardo-Campus 3, 48149 Münster, Germany

[3] Kühne Logistics University;
Großer Grasbrook 17, 20457 Hamburg, Germany

1. INTRODUCTION

New technologies can capture, store, and process vast data [1], spanning millions of terabytes [2], that also tend to be characterized by high volume, variety, and velocity [3]. Due to immense increases in computing power [4], analyses of these big data, or big data analytics (BDA), have become possible, with disruptive and transformative effects [3], taking the form of new business models and ways to conduct business, as well as improved processes for supply chain management (SCM) and logistics [5, 6].

In such settings, BDA is also referred to as supply chain analytics (SCA) [2], which can significantly improve SCM by enhancing existing processes or creating new products or services [7], which has evoked growing interest among both researchers and industry [8]. These realizable benefits are key motives for organizations to implement SCA [9]. Yet they also confront relevant challenges to adopt SCA successfully, reflecting organizational and technical issues such as insufficient resources, privacy and security concerns, data quality, and inappropriate techniques and procedures [10]. Even before they reach these considerations though, organizations must identify applications to which they want to apply SCA and the tools and methods they should adopt for each application. As Schoenherr and Speier-Pero [11] note, no viable guidance exists regarding which supply chain questions can be addressed by SCA.

In particular, prior research offers specific use cases or literature reviews (e.g., [10] for big data analytics capabilities, [12] for applications and techniques), but to the best of our knowledge, there are no publications that seek to analyze SCA from both scientific and practitioner perspectives. Even if researchers might tend to outpace practice, it is important to identify frequently discussed topics in both domains and the gaps between them, which can clarify both application options and areas that demand further research. In attempting to characterize the scientific and practical perspectives on SCA, identify deviations between them, and derive further research opportunities, we seek answers to three main research questions:

1. Which SCA goals and applications are covered by extant literature, and which data and analytics types do they require?
2. Which SCA goals and applications are relevant to practice, and which data and analytics types do they require?
3. How can the gaps between scientific and practical perspectives be addressed by further research?

To answer the first research question and develop a holistic picture of SCA in various literature domains, we conduct a multi-vocal literature review (MLR) [13], which provides a structured identification and analysis of both relevant scientific publications, or white literature (WL), and practice-oriented or grey literature (GL), such as white papers and reports. Here, we offer an overview of recent research, which frames the analytics types and application cases in use, and outline the main motivations for implementing SCA by using particular types of data. The second research question prompts us to conduct a survey among industry experts, who offer their view of SCA and its most relevant aspects, as well as the extent to which they have adopted SCA. The survey results provide a comprehensive view of SCA in practice, reflecting industry actors' main interests, perceptions of relevance, and extent of SCA adoption. Comparing the results pertaining to the first and second research questions inform our insights into the third research question, because it reveals where the two perspectives are aligned, and where they are discrepant when it comes to the applied analytics, SCM tasks, and data they prioritize. The discrepancies also suggest key topics for further research.

Figure 1 depicts the research design. In section 2, we offer some background on SCA and derive the categories we use to classify our MLR and survey results. Then section 3, focused on the scientific perspective on SCA, outlines the MLR methodology and results obtained. Section 4, involving the practitioner perspective, describes the survey and its results obtained. Section 5 provides the comparative analysis, along with the research opportunities it reveals, and then section 6 summarizes the main results, limitations, and outlooks for continued research.

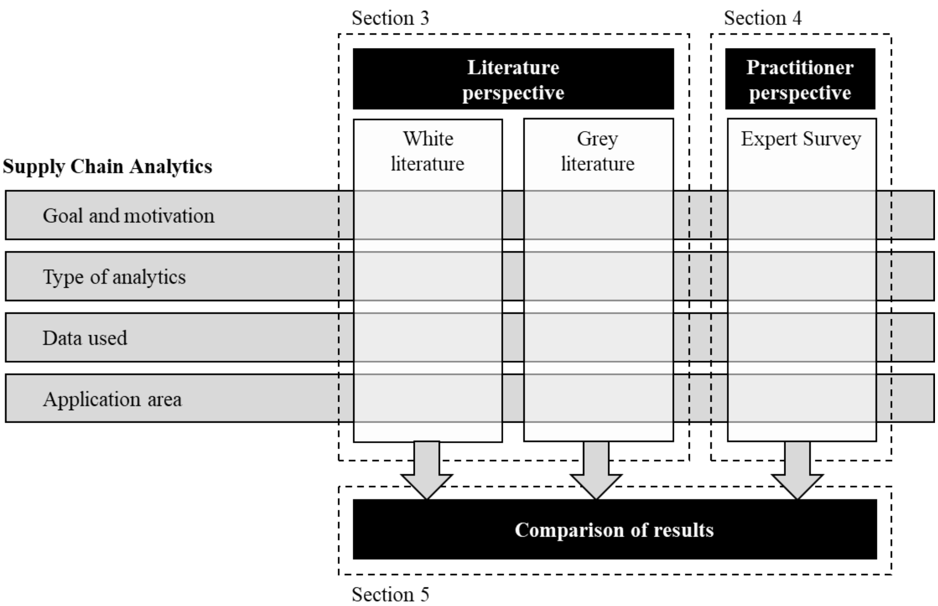

Figure 1: Research Design and Outline

2. BACKGROUND: SUPPLY CHAIN ANALYTICS

To leverage the vast amounts of data they collect, companies increasingly turn to SCA, which encompass qualitative and quantitative techniques to find optimization opportunities in supply chains. They seek to learn from data about existing processes and prior events to shape future decisions. In the next subsection, we will introduce some key definitions and delineate our research from related works. In subsection 2.2 to 2.4, we will also introduce further theories in which we present the key dimensions of SCA that are also used later to classify our findings.

2.1. Definitions and Prior Literature

Table 1 contains a sampling of the many proposed definitions of SCA, which reveals some consistent elements, such as the goal and motivation of the application, the analytics type, the data employed, and the SCM application area for the use case. By integrating these definitions from prior literature we developed a more holistic understanding of SCA as an IT-enabled capability [12, 14, 15] that relies on supply chain data [15, 16] and analytical tools [17–19], leveraged in a particular SCM application area, and is motivated by diverse underlying goals [3, 14, 20]. This understanding will guide our investigation and present the main themes of closer scrutinization of SCA elements.

Table 1: SCA Definitions

Source	Definition
Barbosa *et al.* [21]	"SCA, being Analytics applied to the supply chain, aims at extracting and generating meaningful information for decision makers in the enterprise from the enormous amounts of data generated and captured by supply chain systems."
Cadavid *et al.* [4]	"In the past few years, BDA has proven to be a true advantage for decision support systems, encouraging SC managers to employ these new techniques in their chain. The use of advanced BDA in a Supply Chain, also called Supply Chain Analytics (SCA), encompasses three main branches: Descriptive analytics, Predictive analytics, and Prescriptive analytics."
Chae *et al.* [8]	"Therefore, SCA is viewed as a combination of IT-enabled resources for manufacturing-related data management, supply chain planning and data-driven process and quality improvement."
Chae & Olson [22]	"Drawing upon the dynamic-capabilities literature, this framework describes SCA as IT-enabled, analytical dynamic capabilities composed of DMC [Data Management Capability], APC [Analytical Supply Chain Process Capability], and SPC [Supply Chain Performance Management Capability]."
Min *et al.* [23]	"[…] supply chain analytics refers to a broad range of analytical tools […] that is designed to aid supply chain (SC) professionals in improving supply chain visibility […], identifying specific areas […] of supply chain activities to be improved, and diagnosing supply chain performance outcomes by creating actionable business intelligence which helps SC professionals make more informed decisions."
Tiwari *et al.* [2]	"The term supply chain analytics can be used to define the advanced big data analytics in supply chain management"
Zhu *et al.* [24]	"SCA refers to a group of approaches, organizational procedures and tools used in combination with one another to collect information, analyze information, and gain insights to solve problems and improve performance in supply chain management"
Zhu *et al.* [25]	"SCA refers to the use of analytical tools and applications to inform decision making and ultimately improve levels of supply chain performance"

Previous studies that address applications of BDA to SCM do not cover the range of dimensions that we seek to investigate, as Table 2 suggests. Existing literature reviews, which might involve either WL or GL, focus on specific themes. Some studies turn to existing models to classify results (i.e., deductive), whereas other define categories uniquely for their research effort (i.e., inductive approach).

Table 2: Existing Literature Reviews

Source	Include Grey Literature?	Focus	Categories
Arunachalam *et al.* [10]	No	Capabilities to extract value from big data	Inductively defined
Arya *et al.* [5]	No	Spare parts supply chain of the army	Deductively defined
Aryal *et al.* [26]	No	BDA and Internet of Things, research themes	Inductively defined (semantic analysis)
Awwad *et al.* [27]	No	Challenges and Benefits of SCA	Deductively defined
Chakroun *et al.* [28]	No	Warehouse management	Inductively defined
Chehbi-Gamoura *et al.* [9]	No	Taxonomy for BDA in SCM	Supply Chain Operations Reference model & inductively defined
Moufaddal *et al.* [29]	No	Technological	Supply Chain Operations Reference model
Nguyen *et al.* [30]	No	SCA applications (SCM functions, analytics models and techniques)	Deductively defined
Tiwari *et al.* [2]	Yes	BDA in SCM, no comparison of research and practice	Inductively defined
Fosso Wamba and Akter [31]	No	Development of a research agenda for BDA in SCM	None (in-depth analysis of identified single papers)
Wang *et al.* [12]	No	SCA maturity, strategic and operational applications	Inductively defined

Some existing reviews focus on specific types of supply chains, such as those for spare parts [5], or specific functions, such as warehousing [28]. Other analyses center on certain aspects, such as the challenges and benefits of SCA [27], the capabilities required to extract value from big data [10], technological questions in general [29], or specific technologies such as the Internet of things [26], or aim at developing a research agenda [31]. In contrast, with a more comprehensive approach, Wang *et al.* [12] identify strategic sourcing, supply chain network design, product design and development, demand planning, procurement, production, inventory, and logistics as main application fields for SCA. However, they derive these categories inductively, using scientific research, so they recommend integrating practitioners' perspectives in continued studies. Although Tiwari *et al.* [2] cite GL sources, they do not discuss their differences with scientific literature or outline any gaps between science and practice. In their literature reviews, Chehbi-Gamoura *et al.* [9] and Nguyen *et al.* [30] provide overviews of SCA application areas and widely used analytics methods. In particular, Nguyen *et al.* [30] discuss SCM functions, the levels of analytics (descriptive, predictive, prescriptive), analytics models, and techniques, whereas Chehbi-Gamoura *et al.* [9] propose a taxonomy of BDA applications in SCM, derive a supply chain operation reference (SCOR)-BDA matrix, investigate how SCA is applied, and identify research gaps. However, both reviews stop at 2017, but since then more relevant papers have been published. Moreover, they rely solely on scientific research; Chehbi-Gamoura *et al.* [9] only include journal articles.

Thus, we lack a clear practice perspective and insights into where research and practice deviate. Because GL tends to be published by larger companies or consultancies, which prefer success stories, these publications might not reflect an overall industry view accurately. Nonetheless, by going beyond existing reviews, we provide an updated assessment of SCA applications in scientific research, incorporate GL, and perform a deep dive into practitioners' views with our survey. To guide this investigation, we turn to core elements of SCA, as identified previously.

2.2. Goals and Motivation

Various drivers contribute to the implementation of a SCA use case. Previous studies have established *goals* and *motivations* that are factors in the decision to implement data analytics in corporations [32–34]. In contrast with proposals that key drivers for technology implementation stem from exogenous elements that force technology usage [35], we consider internal factors, too. According to diffusion of innovation theory [36], several steps precede adoption of an innovation, including problem or need recognition [36]. Kwon and Zmud [37] identify technology-based push and problem-based pull effects, which also can be differentiated according to the underlying goals ("What needs to be achieved?") and motivations ("Why is this project being advanced?"). With this nuanced view, it is possible to identify manifold reasons for adoption that might not re flect objective appraisals of the rational aims of an SCA project, as in [32]. To get a glimpse into these reasons, we consider a range of drivers in the category of *goals* and *motivations*. These relate to our other categories as e.g. different *motivations* are associated with varying *analytics types* or *application areas*.

2.3. Analytics Types

In general, analytics types are descriptive, predictive, or prescriptive [38]. These three categories also apply to SCA [10, 20, 39]. Table 3 depicts the three categories and selected methods and tools. As there is no exhaustive collection of methods, the table contains selected examples from multiple review papers. Therefore, the list is not exhaustive nor mutually exclusive, e.g., mathematical programming is a means of optimization. *Descriptive analytics* performs analyses of historical data to understand events that already happened, reflecting the field of unsupervised learning [40], such that the outcome variable is not known. It tries to answer questions about what happened and why [41], by identifying patterns within the data. Common methods include visualization, statistical analysis, outlier detection, or cluster analysis. *Predictive analytics* instead uses historical data to make predictions about trends and events in the future, such that it has a proactive rather than reactionary nature. It aims to answer questions about what will happen, using supervised learning methods to predict a discrete (classification) or continuous (regression) outcome variable [40]. In turn, predictive analytics draws on many different methods, such as time series analysis, regression analysis, and machine learning [41]. Finally, *prescriptive analytics* builds on descriptive and predictive analytics and determines what should happen [20]. Here, optimization is used to assess the impact of specific decisions and choose optimal actions [42]. For SCA, cost reductions through process optimization [43] (e.g., optimal distribution center location) and minimal risks (e.g., simulation of risk occurrences) in the supply chain can be achieved by prescriptive SCA [44]. Methodologically, it relies on optimization (e.g., game theory, multicriteria decision-making such as analytic hierarchy process, optimization techniques) [20], and it frequently turns to simulations to optimize stochastic models [45].

Table 3: Systematization of SCA Components

Type	Selected methods
Descriptive analytics	Visualization [20, 46–49], regression [46, 48], charts and dashboards [46, 49], mapping [20], "pattering" [47], data modeling [48], reporting, online analytical processing [49]
Predictive analytics	Regression [20, 48, 49], time series analysis [20, 47], predictive modeling [46, 48], data mining [20, 48], predictive machine learning, neural networks [47, 49], scoring [46]
Prescriptive analytics	Optimization [46–49], simulation [46, 48], mathematical programming [20, 49], analytic hierarchy process, game theory [20], numerical modeling [46], reinforcement learning [50]

2.4. SCM Application Areas

Thonemann & Papier [51] differentiate between functionally-internal supply chain management (SCM), company-wide SCM, and cross-company SCM. While the primary focus of the paper is the processes in the focal company and thus category two of Thonemann & Papier's classification, we also consider inter-organizational functions of supply chain management, such as the data exchange considerations between organizations elaborated in 4.2.3 and 5.3. For company-wide SCM, the supply chain operation reference (SCOR) model [52] helps to structure applications of SCA according to supply chain processes [9, 29]; it is particularly well suited to provide a structure for SCM tasks and is well-known and widely used in both research and practice. Designed to analyze and measure the performance of a supply chain, the SCOR model has been continuously refined, but it comprises macroscopic supply chain processes: plan, source, make, deliver, return, and enable. These processes are subdivided into several sub-processes. As a benchmark of supply chain performance, its original goal was process improvement

[53], but its broad applicability and substantial degree of detail has made it a quasi-standard for modeling supply chain processes [54]. “Source” deals with the procurement of raw materials and materials needed for production. “Make” refers to the production process, and “deliver” defines the delivery process of the finished goods. Then “return” describes the backhaul process of a product. “Enable” is the process within the supply chain that monitors and controls the entire process. The “plan” process on the other hand includes planning the entire supply chain, including material requirements, finances, and distribution [52]. While this division provides the possibility to classify planning tasks in more detail, it does not provide information on the planning horizon. Hence, the supply chain planning matrix (SCPM) [55] is selected to provide a more detailed categorization of SCOR planning tasks. In proposing the SCPM, Fleischmann *et al.* [55] arrange planning tasks horizontally according to processes and vertically according to the time horizon.

On the horizontal level, planning processes and their planning tasks are divided into procurement, production, distribution, and sales. In our analyses in section 3, we add a ‘cross-functional’ dimension to facilitate a description of those processes that span over more than one mentioned task. In contrast with the SCOR model, the planning process “deliver” is called distribution, and there is a dedicated sales function but no return process. On the vertical axis, the matrix distinguishes planning levels as long-term, mid-term, and short-term. Decisions at the long-term planning level have strategic importance, such as the design or structure of the supply chain. Within the strategic framework, the defined conditions for the mid-term level include, for example, material requirements planning for a month. Finally, the short-term level pertains to operational deployments, such as personnel planning for a specific day [55]. The matrix representation thus establishes a clear listing of all planning components and enables more transparent planning processes [55].

We propose combining the SCOR framework with the SCPM as a way to detail the classification of planning tasks to gain a more holistic overview of relevant supply chain areas for SCA. The relatively sparse consideration of planning tasks (essential application areas for SCA) in the SCOR framework is complemented by the detailed approach of the SCPM. Figure 2 shows graphically how SCOR and the SCPM have been combined and depicts the classification frame for our literature review regarding SCM application areas.

Indeed, SCOR already provides a more detailed categorization of planning tasks. On the second level, plan is subdivided into plan supply chain, plan source, plan make, plan deliver and plan return [52]. However,

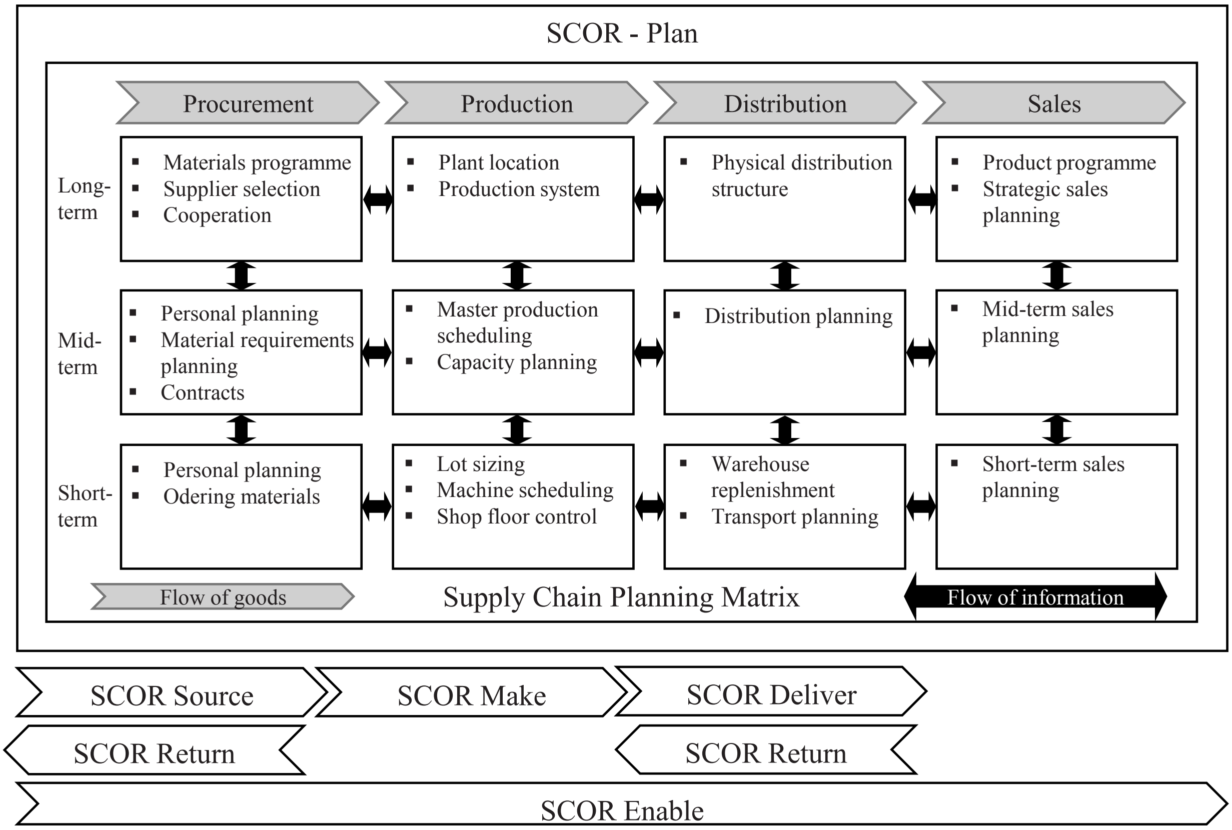

Figure 2: Combination of SCOR [49] and the Supply Chain Planning Matrix (adapted from [51])

this distinction does not provide information about the planning horizon. As planning is one of the major application areas for SCA, distinguishing not only among different planning functions but also time horizons can lead to additional interesting insights. The use of SCPM as a more detailed categorization for planning tasks can offer this benefit. Moreover, the SCPM provides the benefit of a dedicated sales category. SCOR processes do not allow for a categorization of tasks such as sales planning or demand forecasting except for the plan process, resp. the overarching plan supply chain process on a more detailed level. The SCPM enables a distinction of sales planning and other overarching supply chain planning tasks, which would not be possible with SCOR alone. Hence, we chose the SCPM as a detailed categorization over the SCOR plan processes on the second level to classify the literature results described in the next section. However, we did not use this categorization in a later survey study for comprehensibility reasons[1].

3. MULTIVOCAL LITERATURE REVIEW

With a MLR, we seek to identify not just those SCA use cases available in rigorous, published, scientific literature [13] but also include GL, or "publicly available, foreign or domestic, open source information that is usually available only through special channels and may not enter normal channels or systems of publication, distribution, bibliographic control, or acquisition by book sellers or subscription agents" [56]. Because GL generally is written by consultancies or technology providers, it can work as a rough indicator of practical relevance and adoption, though these publications often involve bigger, more mature companies. Moreover, GL is not fully scientific, nor does it claim to represent the state-of-practice. Thus, it offers a limited view of practice, yet it remains a science-focused form of practice literature.

3.1. Methodology

The decision to conduct a MLR must take place in the planning stages of a review; we used a decision aid developed by Benzies *et al.* [56] and adapted by Garousi *et al.* [13] to confirm that we should include GL. The aid consists of seven questions; any affirmative response to any question suggests the inclusion of GL. In our analysis, six questions evoked positive answers (see Appendix A), which strongly supported our use of MLR. Hence, we applied Garousi *et al.*'s [13] methodology to conduct the review. They decompose MLR into three steps: planning, conducting, and reporting [13]. The first step entails establishing the need for MLR, its goal, and the study research questions, as detailed previously. We present the second step in Figure 3.

The research questions provide the input for determining the search keywords. To identify uses of SCA in research and practice, these keywords refer to "supply chain" and "data analytics" and their synonyms. In September 2020, we gathered scientific research from the Scopus and Web of Science databases, which represent the largest bibliographic databases available [57]. We also rely on IEEE Xplore, which specifically contains high-quality technology literature. *Table 4* lists the keywords and query results. The analyzed literature thus spans the past years and is thus focused on a longer timeframe than the empirical data we present in section 4.

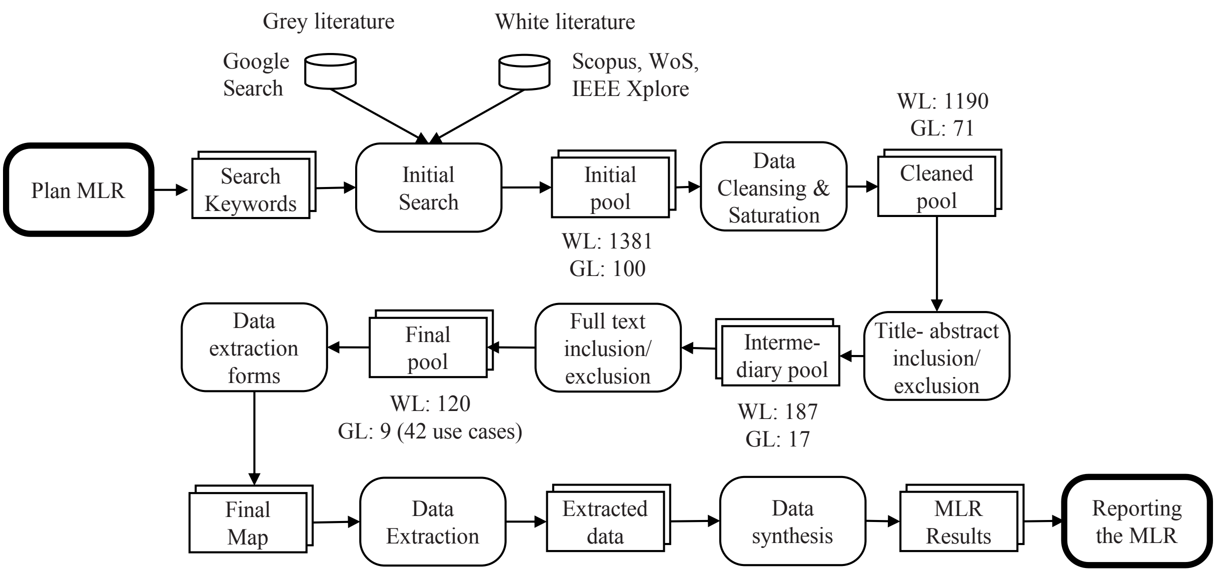

Figure 3: Methodology of this study (adapted from [13])

[1] More details on the reasons can be found in subsection 5.4.

For GL, we queried the Google search engine in January 2021, initially using a keyword string analogous to the one for the WL search. However, this Google search returned almost only scientific, already identified papers. Therefore, we shifted to a much smaller string (cf. Table 4).

Table 4: Keywords for the Literature Search

Database	Keywords	Results
White Literature		
Scopus	TITLE-ABS-KEY (“supply chain analytics” OR ((“supply chain” OR “value chain”) AND (“data analytics” OR “business analytics” OR “descriptive analytics” OR “predictive analytics” OR “prescriptive analytics“)))	710
Web of Science	TOPIC: ((“supply chain analytics” OR ((“supply chain” OR “value chain”) AND (“data analytics” OR “business analytics” OR “descriptive analytics” OR “predictive analytics” OR “prescriptive analytics”)))) Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, ESCI.	536
IEEE Xplore	(“supply chain analytics” OR ((“supply chain” OR “value chain”) AND (“data analytics” OR “business analytics” OR “descriptive analytics” OR “predictive analytics” OR “prescriptive analytics”)))	135
Grey Literature		
Google	“supply chain analytics“ AND filetype:pdf	100

The Google search engine returns massive result sets, so we initially looked at the first 100 results in order of descending relevance.

Overall, we thus reviewed 1381 white and 100 grey literature papers. Reflecting the theoretical saturation threshold proposed by Garousi *et al.*, we reduced the GL set to the first 72 results, as for the last 28 entries, “no new concepts emerge[d] from the search results anymore” [13]. This finding indicates that it likely would not have been fruitful to include more than the first 100 Google search results, and thus, we have not extended the search. We checked for duplicates and erroneous entries. The cleaned pool thus contained 1190 white and 71 grey literature entries. To ensure their relevance to our research goal, we applied the following exclusion criteria:

Formal. Papers could be excluded for formal reasons (e.g., language, entry was an entire conference and not a single document, missing critical metadata). For GL, we also opted to include only first- and second-tier literature, as defined by Garousi et al. [13].

Analytics. Some papers do not discuss a specific analytics method but instead note general frameworks or infrastructure. For example, Hausladen and Schosser [58] deal with data analytics for airline network planning. However, they do not discuss the application of an analytics method but develop a maturity model. Other authors, such as Yanamandra [59], consider a broader set of technologies and were therefore not considered relevant for an analytics-focused paper.

Domain. An article might not use analytics to solve a SCM topic, and the topic could not be assigned to any process category of the SCOR model or any field of the SCPM. Various sources used data analytics to address use cases not specifically restricted to SC tasks but relevant to any company, e.g. human resource management [60] or focused on specific industries such as healthcare [61].

Meta-analysis. Many articles examine the benefits or challenges of supply chain analytics, conduct empirical analyses, or provide reviews. The latter are excluded, because they might lead to duplicates, and the search strings likely do not correspond to the intended scope of this work. This criterion is of special importance for this review. As the goal is to specifically look at the application of SCA, it is necessary to exclude conceptual work or reviews from our result set. Indeed, this leads to the exclusion of highly relevant work such as Fosso Wamba and Akter [62] or Chae *et al.* [63]. While those sources and many other authors provide valuable contributions to the field of SCA research, they do not discuss specific application cases.

Hence, they do not fit to our intended focus on how and where SCA is applied and cannot be categorized according to our chosen frame. Consequently, this review does not cover all relevant work related to SCA in SCM but only those papers with an application focus. Only by excluding conceptual and review works, the later comparison of literature and practice perspectives becomes possible.

Because Garousi *et al*. [13] offer some ambiguity regarding the exact inclusion/exclusion criteria, we adopt a methodology proposed by Thomé *et al*. [64] and analyze the title/abstract and full text separately for inclusion and exclusion. They also propose that exclusion checks be performed by multiple reviewers, with regular reliability checks. For the titles and abstracts, at least two reviewers made decisions independently. Their intercoder reliability rate, calculated as the number of agreed inclusions divided by the sum of all inclusions, was 93.32%. Mismatches were discussed among the group of all reviewers, and all conflicts were solved by mutual agreement. This initial inclusion analysis, using just the title and abstract, led to a pool of 187 white and 17 grey literature papers. Finally, it was reduced even more based on the review of the full texts of the remaining papers. The final pool comprised 120 WL and 9 GL publications. Despite the difference in these values, the latter reports each include numerous use cases, involving multiple analytics applications and use cases (42 total), whereas most scientific papers contain only single use cases. Thus, the actual extracted data are roughly on par.

With this final pool of papers, we conduct analyses to gather valuable information. To retrieve data in a structured manner, *Garousi et al*. [13] suggest designing extraction forms first. On the basis of our theoretical background (section 2), the use of SCA entails multiple dimensions:

Goal and Motivation. We turn to prior studies [32–34] to generate potential *goals*: “optimize internal processes,” “optimize decision,” “optimize existing products/services,” “develop new products/services,” “acquire new customers,” and “capture and utilize new knowledge.” The identified *motivations* are “utilize existing data,” “solve a specific problem,” “use new technological possibilities to communicate this internally/externally,” “increase transparency for a particular area,” “draw level with competition already using SCA,” and “satisfy the demand by partners for analyses.” With these options, we can establish comparisons with practitioner responses, across broader set of potential replies. Not all options are relevant for the analysis of this literature set alone.

Analytics. We recorded the applied analytics methods of each paper and assigned the type of analytics to descriptive, predictive, or prescriptive, as defined in Table 3 of section 2.

Data. The input data for the data analytics methods were subsumed under more general data types when possible, without compromising specificity too much (e.g., Facebook data included in social media data).

SCOR/SCPM. The papers were assigned to one or multiple of the major processes of the SCOR model [52]. For the plan and enable processes, they also were assigned to a specific process category, if possible. As described in subsection 2.4, the SCPM [55] serves the purpose of further detailing the SCOR plan process. Thus, all papers assigned to SCOR plan have also been assigned to one category from the SCPM.

3.2. Results

Supply chain analytics is an emerging field, and research has grown immensely in recent years, as Figure 4 reveals. Among the 129 publications we consider, most works have been published in the past five years, with 90 publications since 2018. The average annual growth rate is 55.44%, whereas the annual growth rate of all SCM publications [65] is 12.88%, marking SCA as a rapidly emerging topic.

3.2.1. Goals and Motivation

Figure 5 shows the share of use cases that applied SCA for six different goals. We find that WL applies SCA to optimize decisions to the greatest extent, with almost two-thirds (63%) of papers citing this goal. Far behind, the optimization of internal processes appears in nearly one-quarter (23%) of publications, followed by gathering and using previously unused knowledge (11%). The application of SCA to optimize existing or develop new products or services is marginal; acquiring customers seems irrelevant to WL. In the set of GL, the top two spots switch, such that optimizing internal processes is the topic of almost three-quarters of use cases (73%). About every fifth case uses SCA to optimize decision-making (19%), and every tenth gathers and applies untapped knowledge (11%). One of these use cases refers to creating new products and services through SCA [66], but two goals, develop new products/services and acquire new customers, are not addressed at all. In summary, the most notable difference is that in the GL, the focus is on optimizing internal processes, whereas in WL, it is on optimizing decision-making.

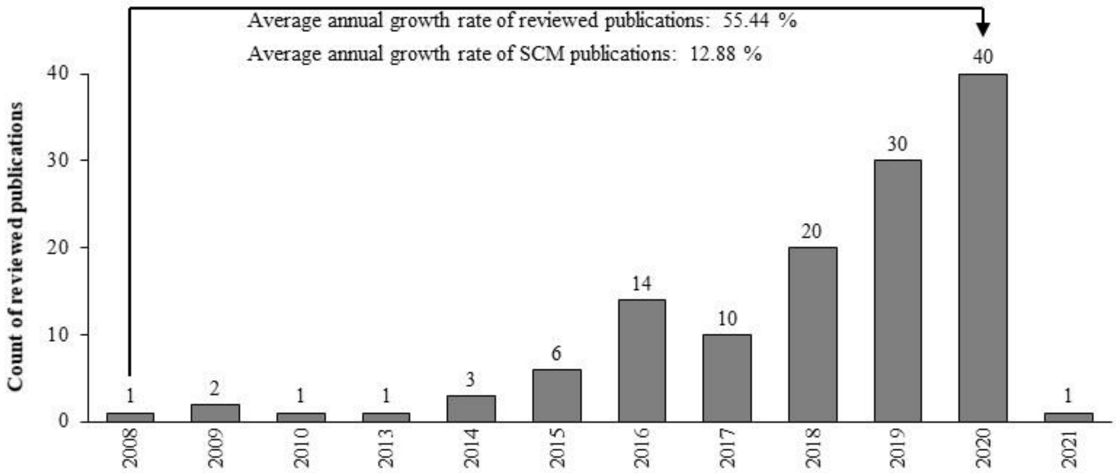

Figure 4: Publications per year

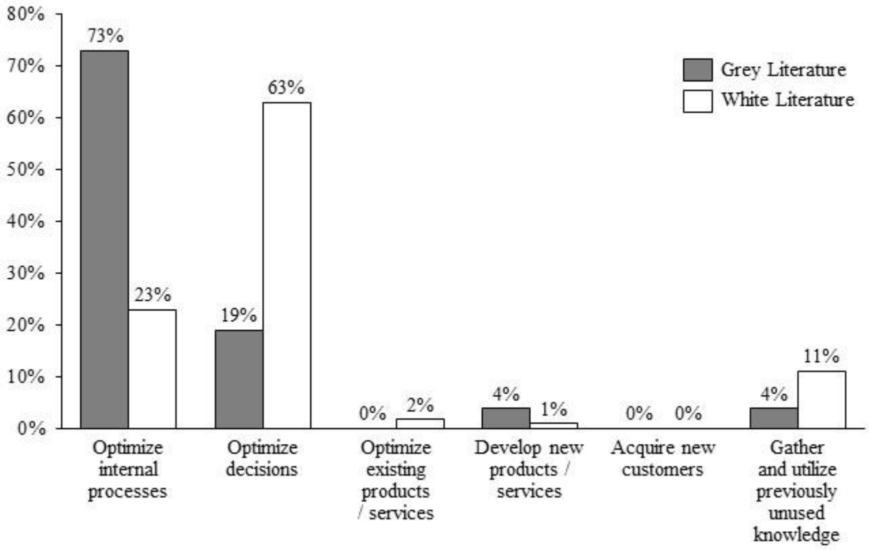

Figure 5: Goals underlying the respective SCA use case

Figure 6 details different motivations for employing SCA and the shares of WL and GL for each. A major WL focus is solving one specific problem, as occurs in more than half of the cases (54%). This motivation is far more prevalent than improving transparency, the motivation underlying every fourth paper, or utilizing available data, which was the motivation for every fifth case. Only one study used SCA to deploy new technologies (for external communication) [67]; the three remaining motivations are not addressed at all in WL. In contrast, GL focuses on achieving increased transparency (55% of use cases). 18% of use cases rely on SCA to utilize available data, and one-tenth of them address the motivations solve a specific problem, deploy new technologies (for external communication), or partners demand the use. Two motivations are not addressed at all. These discrepancies make sense, considering the target audiences of WL and GL. The former seeks progress on the scientific edge of research, so it seems sensible to focus deeply on individual problems to be solved. But GL seeks to provide value for the primary consumers of these reports: the companies. They have strong incentives to generate transparency and offer data at a higher vantage point. The focus on individual solutions for particular problems is secondary. Still, the value-adding potential of offering best practices, based on individual use cases, should not be neglected. Finally, both areas consider the usage of available data as a valid motivation.

3.2.2. Analytics

Figure 7 provides an overview of how many sources discuss descriptive, predictive, and prescriptive application cases. The applications of descriptive

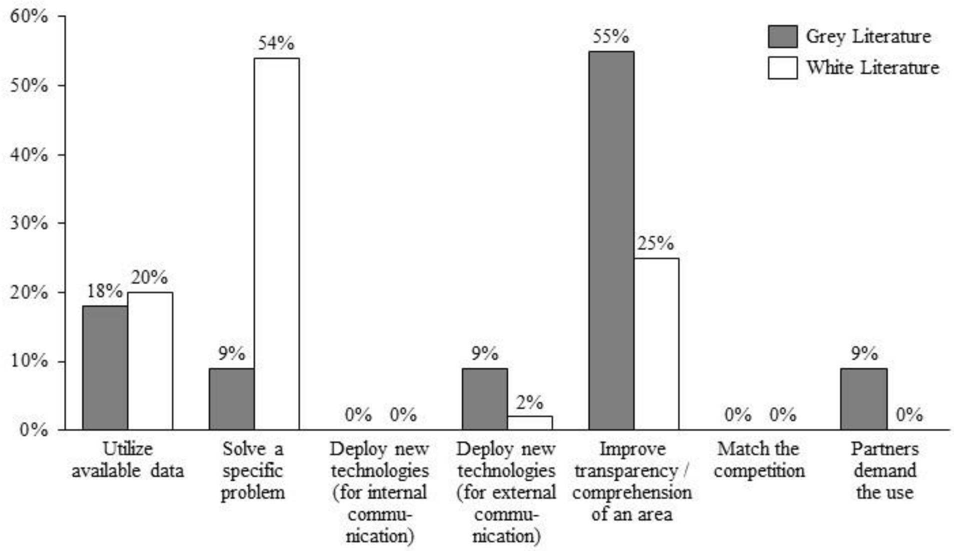

Figure 6: Motivation for investigating the respective SCA application

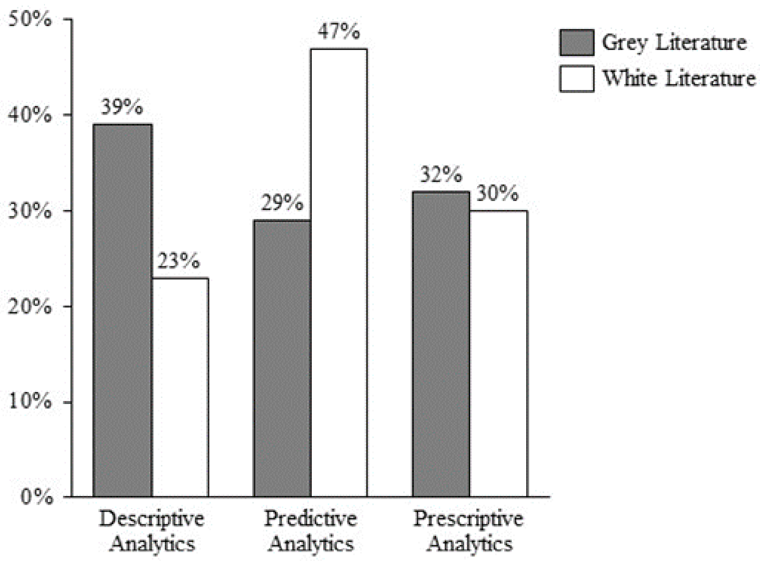

Figure 7: Prevalence of SCA types (literature)

analytics (23% WL, 39% GL) mainly revolve around supply chain transparency and visibility, such as by identifying often-ordered products [68] or implementing a tool that provides real-time visibility into the operations of a pharmaceutical company [69]. Most works used different statistical analyses such as vector autoregression [70] or correlation analyses to investigate the dependence of multiple variables [71]. Predictive analytics (47% WL, 29% GL) entails regression and classification tasks, such as demand forecasting [72, 73]. But regression and classification are also applied in other SC processes, such as machine failure prediction [74] or supplier classification and selection [75]. Classification and regression were performed roughly equally among all publications; however, multiple different algorithms were used for these two applications. For the former, decision trees [76, 77], random forests [78], neural nets [79], or k-nearest-neighbor [68] algorithms are applied. For the latter, simple linear regressions [76], ARIMA [70, 80], or more complex machine learning models, such as support vector machines, neural nets, and many more [81, 82], are used. Finally, prescriptive analytics is the least discussed category (30% WL, 32% GL), and these papers use optimization or simulation to assess the consequences of decisions or test alternative scenarios, such as for determining optimal order quantities [83]; they also might analyze supply chain networks' performance, depending on different market behavior scenarios [84]. For that, mostly optimization methods were used, including mathematical optimization [85–88], heuristics [89, 90], or genetic algorithms [91].

Our analysis reveals that WL is heavily focused on predictive analytics, which accounts for almost half of all included papers. Advancements in machine learning and AI indeed have extended the possibilities of predictive analytics. In contrast, we note a general neglect of descriptive analytics, which might stem from the relative simplicity of this analytics type. Limited scientific interest also seems to apply to prescriptive analytics, even though it is arguably the most mature type. Among GL, we find a more decentralized focus, such that all three types are represented more or less evenly, except for a slight dominance of descriptive analytics, which is fundamental to the operations of any supply chain, followed by prescriptive analytics, which tend to offer more actionable results.

A comparison across analytics types reveals a notable difference between WL and GL: Predictive analytics are most prominent in scientific literature (47% WL, 29% GL), but practice publications focus instead on descriptive analytics (39% GL, 23% WL), seeking first of all transparency and understanding of the data. In contrast, scientific literature tends to assume the data are available in suitable form to apply predictive analytics. This differentiation might stem from the fact that scientifically utilized data-sets in WL publications have a selection bias towards more sophistication to establish the cutting-edge nature of the research conducted. Comparatively, the considered data-sets in GL can be more representative of the state of data availability and quality in industry, which to a considerable degree do not facilitate predictive and prescriptive approaches. This gap needs to be investigated further.

3.2.3. Data

In line with the distinct analytics foci between WL and GL, the data exploited by these research streams also differ, as Table 5 shows.

Table 5: Data Types in WL vs. GL

Data type	**White Literature**			**Grey Literature**			**Difference**
	Rank	**#**	**%**	**Rank**	**#**	**%**	
Transport data	1	26	10.40%	7-11	5	5.95%	-4.45%
Inventory data	2	23	9.20%	1	10	11.90%	+2.70%
External data	3	23	9.20%	7-11	5	5.95%	-3.25%
Demand data	4	17	6.80%	16-18	1	1.19%	-5.61%
Cost structure data	5	17	6.80%	2	7	8.33%	+1.53%
Product master data	6	14	5.60%	7-11	5	5.95%	+0.35%
Website data	7	13	5.20%	-	0	0.00%	-5.20%
Demand forecasts	8	12	4.80%	3-6	6	7.14%	+2.34%
Production planning data	8	12	4.80%	3-6	6	7.14%	+2.34%
Point-of-sale data	10	11	4.40%	7-11	5	5.95%	+1.55%
SC network data	11-12	10	4.00%	3-6	6	7.14%	+3.14%
Production capacities	11-12	10	4.00%	3-6	6	7.14%	+3.14%
Sales quantities	13	8	3.20%	12-13	4	4.76%	+1.56%
Quality data	14	7	2.80%	12-13	4	4.76%	+1.96%
Sensor data	15	6	2.40%	14	3	3.57%	+1.17%
Purchase order data	16-19	5	2.00%	-	0	0.00%	-2.00%
Material flow disrupt-tion data	16-19	5	2.00%	16-18	1	1.19%	-0.81%

Goods receipt data	16-19	5	2.00%	7-11	5	5.95%	+3.95%
Social media data	16-19	5	2.00%	16-18	1	1.19%	-0.81%
Logs	20	3	1.20%	-	0	0.00%	-1.20%
Other	21-34	17	6.80%	15	2	2.38%	-4.42%
- Performance data	24-34	1	0.40%	15	2	2.38%	+1.98%

Transport data are most frequently used in WL, appearing in 26 papers, or 10% of all data. For example, GPS data support a cluster-based frequent trip analysis to help fleet managers and dispatchers improve estimates of the time of arrival [92]; transit types, product categories, and shipping destinations have been entered into a decision tree that scrutinizes the causes of cargo loss severity [77]; and a mixed-integer nonlinear program with carrier capacities and carbon emissions facilitates sustainable procurement and transportation decisions [86]. This focus on transport data, as well as external or website data, might reflect the relatively easy access they offer. In contrast, inventory data are most used in GL (second for WL), such as in a descriptive materials management analysis designed to quantify the effectiveness of a supply chain's material asset management [93]. A prescriptive analytics tool applied to inventory data also suggests ways to optimize the supply chain [94]. Cost structure data, which are highly confidential, are more frequently addressed by GL.

Although inventory data achieve similar rankings in WL and GL (second and first), other data uses highlight significant deviations. For example, demand and website data were employed less in GL (-6% and -5%). This finding is surprising; demand data are common elements of sales forecasting analytics for example. In addition, the demand data often are combined with other data types in WL. Furthermore, *Boldt et al.* forecast Nike sales using website data such as Google Trends and Facebook data [71]. Another study gathers data from Twitter and Facebook to analyze sentiments and trends, then combines these findings with historical sales to predict product demand [95]. In this sense, perhaps the lack of website (0 papers) and social media (1 paper) data adopted by GL explains the lack of demand data. A thorough analysis of why these data types have not been adopted in practice is required though.

More traditional data sources also exhibit a more significant share in GL than in WL, with a maximal deviation for goods receipts (+4%), supply chain networks, inventory data, and production capacities (+3%). We discern no particular reason for the increased use of these data, other than the aforementioned spikes in research that employs sales or website data.

3.2.4. Application Areas

The upper part of Figure 8 depicts the distribution of SCOR processes in literature. In both WL and GL, we find a strong focus on planning processes (66% vs. 62%), reflecting the great relevance of demand and sales forecasting application cases [81, 96].[2] Enable is the second-most discussed category (28% WL, 14% GL). These studies deal with risk management [79], managing the supply chain network [97], or performance management [98]. In particular, WL seems to assign less priority to issues related to operative supply chain functions. Source, make, deliver, and return processes together appear in only 6% of all WL, whereas GL include at least one of them in 24% of the cases it reports. The most substantial gap pertains to the source process: Only one scientific research paper can be assigned to this category, including a quality assessment method for incoming goods [99]. But the GL set contains various application cases, such as for commodity pricing, identification of delivery mismatches, and material management [69, 74, 93]. Source and the other operative processes signal a slight misalignment between WL and GL. The industry seems to focus on using SCA for operational functions. Perhaps researchers have not matched this focus on operational tasks due to the lack of data sets or a sense that these tasks already have been resolved satisfactorily.

In contrast, we find strong alignment regarding the application potential of SCA for planning processes. Both research and practice identify viable opportunities in this category. Accordingly, a more detailed categorization and discussion of this process seems reasonable. That is, we apply the SCPM to gain

[2] As we focus on scientific literature in this section, the SCOR model is deployed in its common form as a well-established scientific model. However, in subsection 5.4 when comparing our literature findings with empirical results, we interpret the survey respondents by re-classifying the SCOR processes. More details on the reasons for and implementation of this re-classification can be found in subsection 5.4.

insights into which planning tasks are of special interest by assigning each paper in the SCOR plan process to a category of the SCPM as shown in the lower part of Figure 8. This is beneficial as, unlike SCOR, the SCPM provides a dedicated sales planning category, allowing to assign demand forecasting tasks to it.

Focusing on WL, this categorization makes the aforementioned focus on demand forecasting evident: 30% of WL sources focus on sales planning, and 18% deal with mid-term sales planning, mainly related to mid-term demand or sales forecasting, as might be gleaned from Facebook data [71]. However, demand forecasting

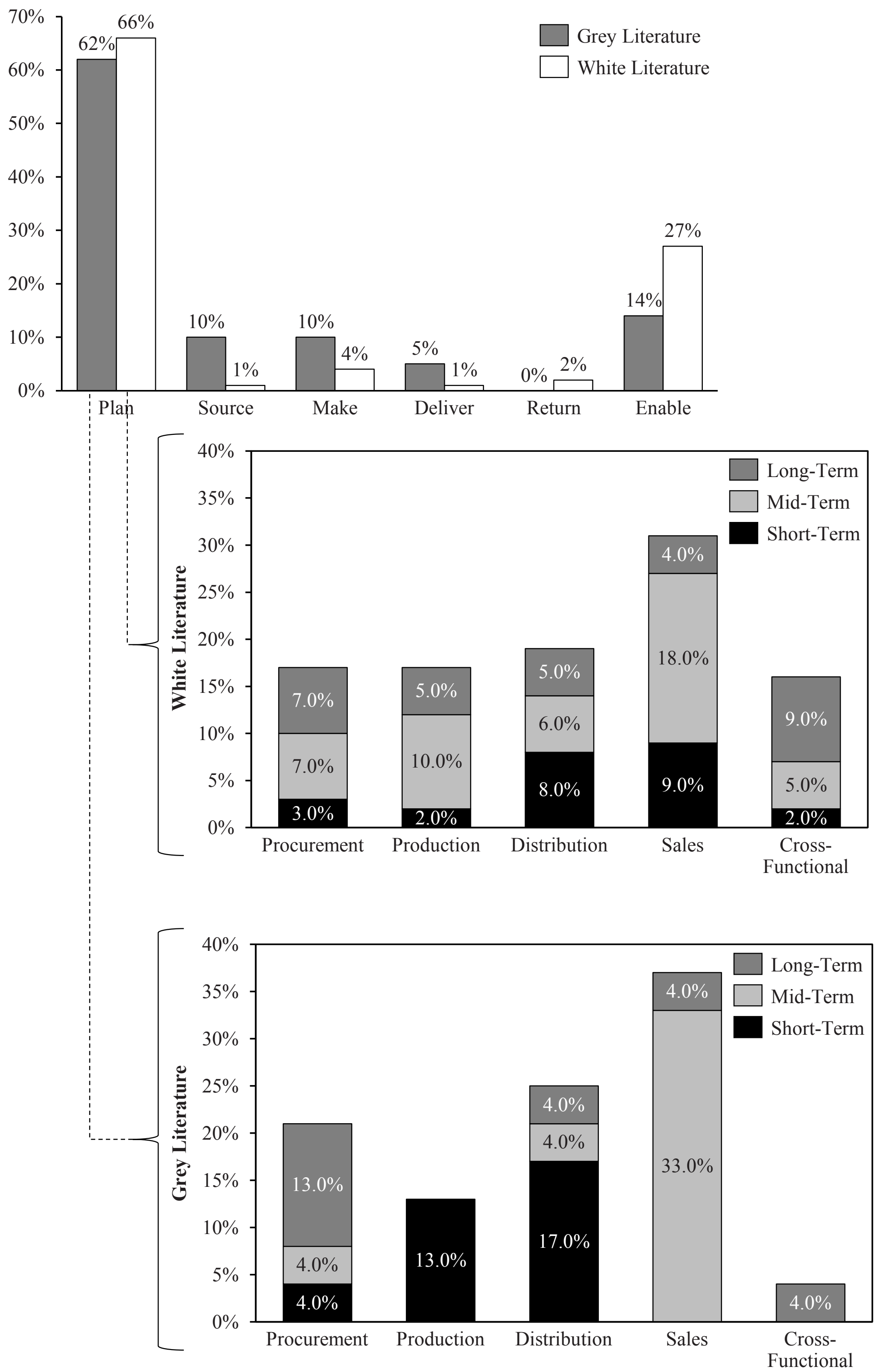

Figure 8: Prevalance of SCOR processes and a focus on planning tasks addressed by literature

also spans the long- and short-terms or combines multiple time horizons. For example, Isikli et al. use consumer reviews for short-term demand forecasting of configurable products [100], and Punia et al. propose a hierarchical framework to generate short-, medium- and long-term forecasts [83]. Otherwise, all the other planning functions are roughly qually represented. Procurement planning is the focus of 18% of studies, about equally split between long- and medium-term considerations. Most sources referring to long-term procurement address supplier selection [79, 101], wherea the mid-, and short-term planning studies revolve around order allocation [86] or inventory-related decisions [90, 102]. Similarly, most sources dealing with production planning address a medium-term time horizon to cover topics such as predictive maintenance [103] or the design of production processes [68]. Distribution planning is slightly more discussed (19%), with a predominant short-term time horizon that focuses on tasks such as route planning or prediction of arrival times [104]. The growth of e-commerce and altered consumer expectations regarding delivery speed make features such as on-time delivery or anticipatory shipping [91] more important, as already reflected in extant literature. Finally, the additional category of cross-functional planning, i.e. planning for multiple functions at once, had to be added to reflect such sources correctly. Cross-functional planning appears in 14% of the sources, for which the importance of long-term planning is evident. Cross-functional planning with a long time horizon tends to deal with network design or location planning, such as designing an international manufacturing network [87] or selecting storage and fulfillment locations, along with strategic supply chain partners [68]. Other issues assigned to this category are more general or strategic in nature, such as collaboration or coordination within a supply chain [97, 105].

Among the set of GL, the distribution of sources is roughly the same across planning functions but deviates in the planning horizons. Again, the prevalence of demand and sales forecasting, and thus sales planning, is obvious (38%). The focus on mid-term sales planning is even stronger than in the WL set. Authors use SCA to attain more accurate forecasts for an auto parts supplier [69] or anticipate demand for certain food products during hurricanes [74]. Regarding procurement planning, we find no remarkable differences with WL: Long-term procurement with a focus on supplier selection or sourcing strategy development is the most addressed topic in GL too [93, 106]. However, with regard to production and distribution, GL focuses on short-term planning tasks, unlike WL, which implies a critical gap between the interests of scientific researchers and practitioners. Apparently, the latter regard possibilities to use SCA for short-term production or distribution planning as more relevant. The publications include efforts to predict arrival times by using GPS, traffic, and weather data [73] or schedule production more accurately by leveraging enhanced data visibility [74]. Finally, we identify another form of variance in the cross-functional planning studies. Whereas WL addresses this category in 14% of cases, GL only includes it in 4%, and always in a long-term planning context. For example, Sandiford mentions the possibility of analyzing supply chain networks and comparing different future scenarios to improve them [84]. However, this example is a rare exception, because GL usually focuses on individual companies instead of overarching networks. The gap in the importance of cross-functional planning might stem from the same reasons we offered for the varying perceived relevance of short-term planning. However, these suppositions and whether they are reflected in industry experts' views require confirmation.

4. EXPERT SURVEY

To evaluate the GL results, regarding practical applications of SCA, as well as seek any other potential deviations between practice and research, we conducted an expert survey.

4.1. Methodology

The empirical basis for this analysis is part of a broader survey on trends and strategies in logistics and supply chain management, conducted in cooperation with the German Logistics Association (BVL), a non-profit association that represents logistics and supply chain management professionals. Thus it represents a wide variety of the primary roles in supply chains, including production, logistics, and trade. An online questionnaire, designed in 2020 and distributed among the BVL network, generated results that provide a representative view of the field. Data were collected from the beginning of February to the beginning of March 2020, so it took place immediately before and during the initial outbreak of the COVID-19 pandemic. This factor may have influenced the survey results and thus should be considered when drawing inferences from our data. The survey was available on the BVL website and sent via its e-mail distribution list, such that it was accessible to any interested members. The approximate response rate, based on the volume of contacted individuals across all channels, is 2.3%, which produced 276 complete responses. We break down the sample by economic sector and company size in Figure 9, demonstrating that small and medium-sized enterprises (per the EU definition) make up 23% of our sample, midcap firms represent 30%, and large companies are 47% of the sample. These demographics match the overall membership structure of the BVL [107, 108], so our sample is a representative foundation of supply chain experts from Germany, Austria, or Switzerland.

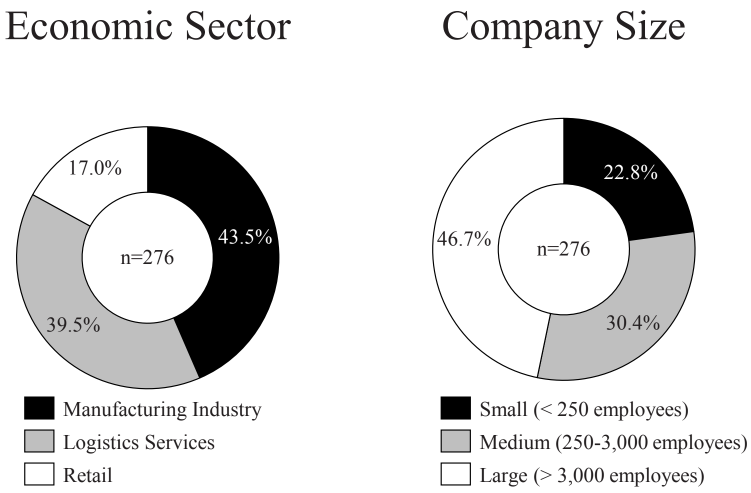

Figure 9: Demographics of the survey

The survey contains a general section, focused on macroscopic trends, strategies, and technologies applicable to the entire industry. All participants responded to it. Then three sections addressed the core areas sustainability, data analytics, and talent management questions in more depth. Each participant was assigned to one core area; surveys related to these topics represent subsets of the overall data set. We gather data from both the general section and the data analytics core area. This is the reason why some of the following analyses are based on different response counts. We consistently state the specific sample size n, to clarify whether the findings reflect the general survey or core area data set.

4.2. Results

In the following, the results of the expert survey analysis are presented before they are juxtaposed to the literature results in the next section. The survey questions and analyses were again separated into the previously presented dimensions “Goals and Motivation”, “Analytics”, “Data”, and “Application areas”.

4.2.1. Goals and Motivation

Similar to the MLR, we analyzed the goals that practitioners strive to achieve with SCA (n = 60). The results in Figure 10 highlight that optimization is the number one goal for practitioners to apply SCA. They ranked optimize internal processes and optimize decisions as high in relevance, whereas develop new products and services and acquire new customers receive moderate ratings. This might be due to the fact that it is more obvious to practicioners to improve exsisting processes or decisions, as they are aware of their shortcomings, than to move to new grounds and develop new prodcuts/services or acquire new customers. In addition, the latter are generally not executed as often as internal processes or decisions. Record and utilize unexploited knowledge and optimize existing products or services receive moderate to high ratings. In addition, the use of SCA is motivated by different reasons in practice (n = 64), as shown in Figure 11. The strongest motives are to use existing data to add value, solve specific problem, or increase transparency/understanding of an area, which supports the impression of the goal analysis, that SCA is rather viewed as a means to improve existing, internal processes or decisions. Again, the extrinsic perspective, i.e. the motivation that partners demand the use or to match the competition is not as relevant as the intrinsic one to use SCA.

4.2.2. Analytics

Noting the different types of analytics, the respondents (n = 299) categorized their relevance; as Figure 12 indicates, descriptive and predictive analytics were both rated as moderate to high in relevance, whereas prescriptive analytics has slightly less relevance for practitioners. For the actual implementation of SCA in practice, we also note that the more sophisticated the analytics, the lower the adoption level (see Figure 13). For example, 60% of respondents (n = 276) have implemented descriptive analytics at least marginally, but only 47% have implemented predictive and 30% prescriptive analytics. In coming years, the gap between descriptive and predictive analytics arguably might close, considering that 28% of respondents stated that they planned to implement the latter. The descriptive analytics implementations are only planned by 18%. The share of respondents that plan to implement prescriptive analytics is higher (30%), but so is the share of practitioners that have not planned any such implementation (39%), likely reflecting the comparatively lower relevance they assign to it. Overall, these numbers reflect the idea, that more complex analytics types are both seen as less relevant and are less implemented. The latter one is easy to explain

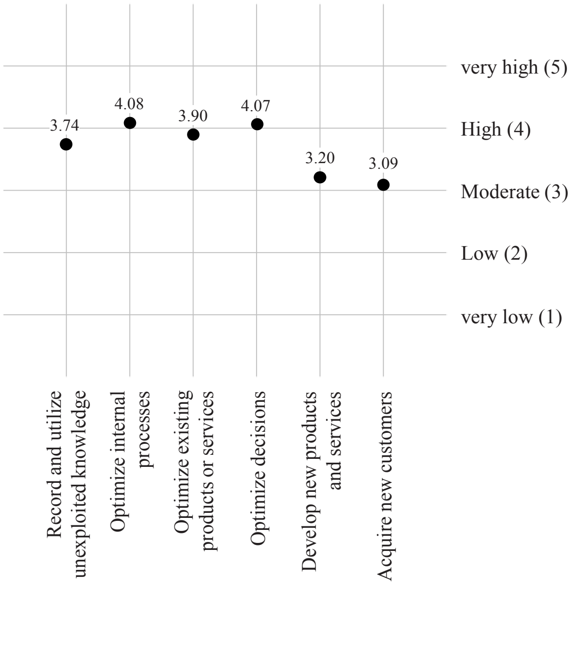

Figure 10: Goals of SCA (survey)

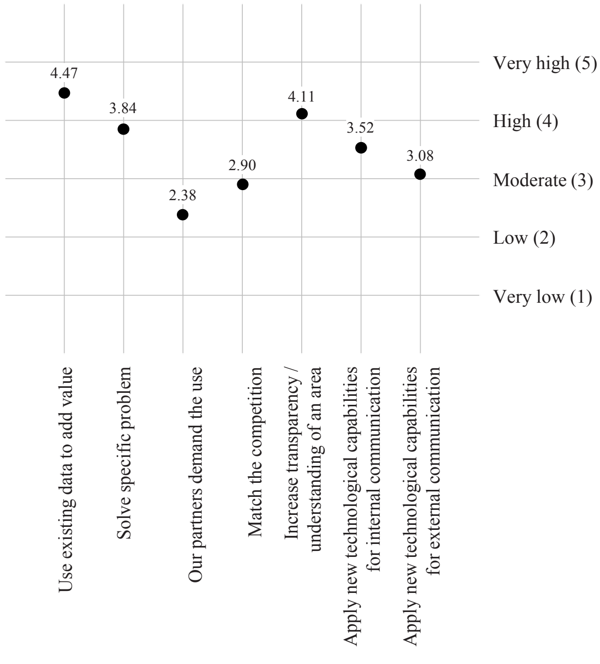

Figure 11: Motivation for SCA (survey)

due to more resources needed for the implementation of complex algorithms etc. Similarly, practitioners might not be able to think of reasonable and promising business cases for more complex analytics types. Hence, they do currently not see the added value they could bring and assign them a lower relevance.

Figure 14 reveals the gap between the relevance of SCA and its implementation in a lollipop plot (n = 299). The lower dots signify the share of respondents who classified the relevance of the different analytics types as "high" or "very high." The upper dots depict the share of respondents who have implemented SCA, partially or extensively. Then the black lines represent the gap. The gap for descriptive analytics is the smallest (21%), such that solutions already are quite mature and applied (relevance = 64%, implementation = 43%). In contrast, the biggest gap involves predictive analytics, such that slightly more respondents think it is highly relevant (66%; cf. descriptive analytics), but their implementations are vastly lagging (25%). Here, practicioners already seem to recognize many application cases and thus assign high relevance to predictive analytics, but have only marginally been able to impelement their ideas. That is, 41% of respondents

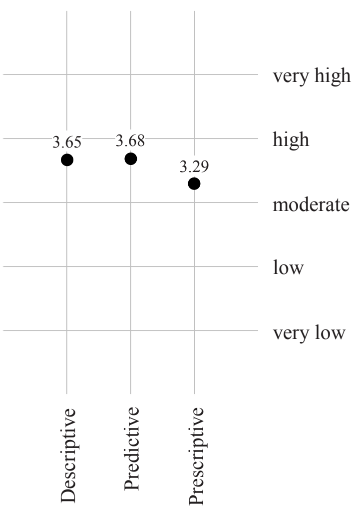

Figure 12: Relevance of SCA Types (survey)

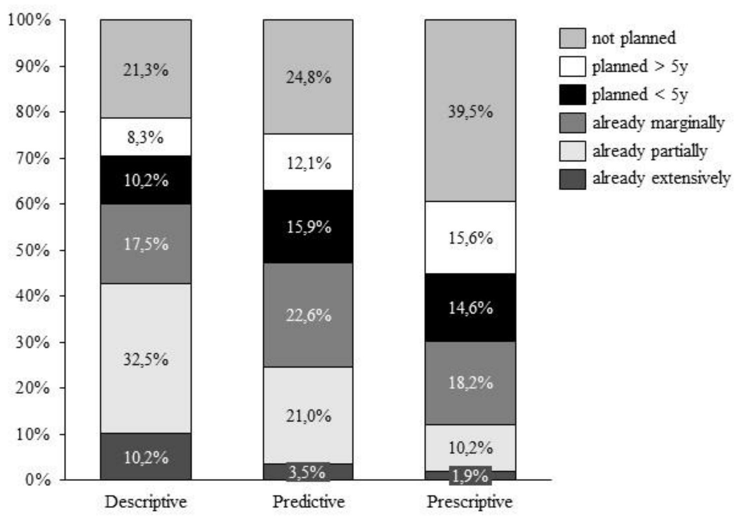

Figure 13: Implementation of SCA (survey)

consider predictive analytics highly relevant but engage in marginal or no implementation. For prescriptive analytics, the level of implementation is lower (12%), but so is its relevance (47%), as practitioners might not foresee suitable use cases for prescriptive analytics, so its 35% gap falls between predictive and descriptive analytics.

4.2.3. Data

Beyond the relevance and implementation of different analytics types, the respondents (n = 79) also indicated which data were available to them to perform supply chain analytics,which data might be missing and which data they share. Figure 15 indicates the data that practitioners marked as needed (upper dots) versus actually available (lower dots) for their supply chain analytics. The y-axis shows the share of all respondents, and the black lines, that reflect the gap, pinpoint where supply chain analytics is well-enabled or else hindered by the (un)availability of particular data. Overall, the average gap between missing and available data is 34% (standard deviation: 8%). Transport data are desperately needed, by 87% of respondents, and also the most available data, at 59% (willingness to share: 57%). Goods receipt and product master data also indicate small gaps.

Arguably, all these data are rather easy to collect and get for a company, as they are purely internal data that neither need collaboration with supply chain partners nor the application of any algorithms to be available. In contrast, demand forecasts and quality data are strongly required (82% and 78%) but not widely available (35% and 38%), possibly because its rather difficult to acquire reliable forecasts or to decide on measures for quality. The willingness of firms to share these data is also equally low (38% and 39%). These gaps between planned and already performed SCA indicate what data types are missing and what might be hindering the implementation of future analytics initiatives. Another big gap appears in relation to material flow disruption data, available to less than one-quarter of respondents (20%). These data seems to be especially hard to collect as monitoring of large parts of the supply chain would be required. Considering that this gap exceeds 51%, we note vast potential for improvement, relative to the other data types that indicate a more moderate gap around 31% (standard deviation: 4%). Unfortunately, the willingness to share these data is also at 20%. Lastly, the availability of R&D data is the lowest (9%), and even fewer firms want to share (6%), as they represent one of the most confidential types of data. Ultimately, the successful adoption of SCA in practice requires closing these gaps and making more data available, but the reasons ranging from the inability to collect to the unwillingness to share the data are manifold.

4.2.4. Application Areas

We classify how respondents (n = 67) rate the potential of SCA for different supply chain management areas using the SCOR model. Most respondents identify strong potential of SCA for supply chain planning (average 4.38) and a high relevance for delivering (3.96) and sourcing (3.74). For make processes, we find a moderate rating (3.44), but return appears to be the least promising process category for SCA, with a score of 3.25. To assess the potential of the different analytics types for the SCOR process categories, as depicted in Figure 16 (n = 67), we count the number of respondents who rated the relevance of the different SCA types and the potential of SCA for SCOR as either "high" or "very high." Generally, plan is the highest-rated category (91

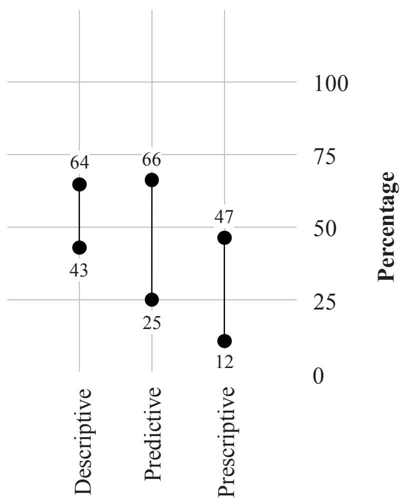

Figure 14: Gap between relevance and implementation of SCA (survey)

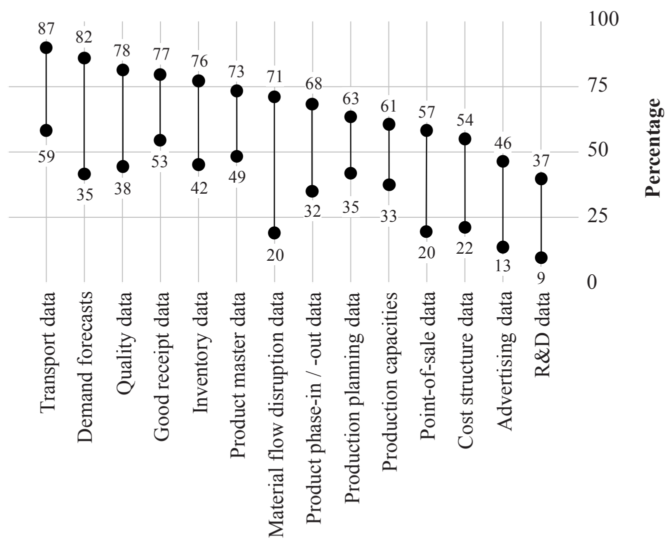

Figure 15: Needed data, available data and the gap (survey)

responses), closely followed by deliver (88) and source (76). Make (54) and return (46) fall to the bottom. The numbers show that practitioners see the biggest potential for SCA in plan, deliver and source processes, which hints at current struggles in these categories. Possibly, the need to consider more than the own company for those processes leads to practitioners constituting high improvement potential. The 'enable' category of the SCOR process was not included in the survey due to the indirect nature of the category and the subsequent need to elaborate on the underlying processes. In pre-tests of the survey we realized that the category is not as self-explanetory as the others and the participants are at risk to have a different understanding of the nature of the category. Due to the need for brevity in the survey we decided that it was not advisable to try to provide information to generate a more coherent picture of the category among participants and did not include it in the survey.

The intersection between the SCOR model and SCA types reveals that proportionally fewer respondents acknowledge the potential benefits of descriptive analytics for make processes (33.33% of responses) than for plan processes (36% of respondents). Yet for return processes, prescriptive analytics are identified as high-potential options by 28% of the participants, and for the plan process, this percentage drops to 24%. Finally, predictive analytics really shines for the category deliver (41%), not for return processes (37%). Interestingly, there is no clear tendency to which SCOR process are regarded as suitable for which analytics types. This was also supported by a Pearson correlation analysis which showed the highest correlation between prescriptive analytics and make processes with a weak correlation coefficient of 0.30 ($p = 0.03$). Other combinations led to even lower correlations and p-values above 5%. Again and for all processes, prescriptive is considered less relevant.

5. COMPARATIVE ANALYSIS

The comparison of perspectives represented in scientific literature and practice can uncover similarities and differences, which also suggest research opportunities. In addition to consistent sections dealing with goals and motivation, analytics types, data, and application areas, we present a cross-analysis of application areas and analytics types at the end of this section.

5.1. Goals and Motivation

The relevance of different goals for implementing SCA in practice corresponds somewhat to the MLR findings. The goals of 'optimiz[ing] internal processes' and 'optimiz[ing] decisions', which are focused upon in many publications, also achieve a high average relevance rating from survey participants. The motivations to use existing data to add value, solve specific problems, and increase transparency/ understanding of an area, predominant in the MLR, also attain the highest average relevance rating across all goals (high) in the survey. These goals and motivations are directly attainable with SCA and therefore predominant in literature and rated higher. In contrast, more indirect goals, like 'develop new products and services' and 'acquire new customers', instead receive moderate ratings; they similarly were not widely represented in the literature. This may be

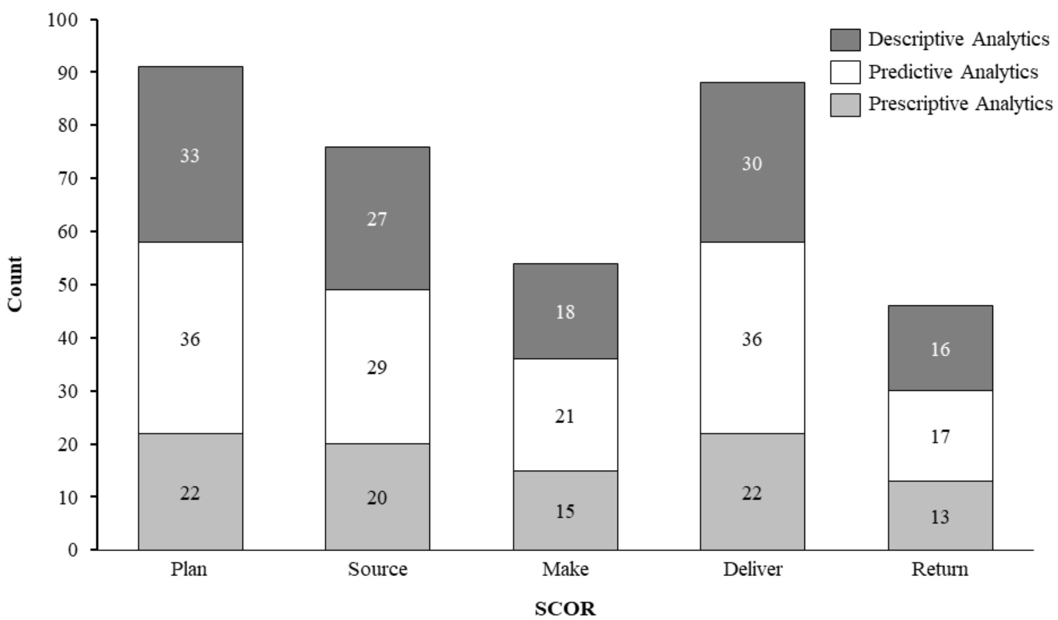

Figure 16: Potential of SCA types per SCOR process

the case due to the more indirect nature of adding value when focused on these goals – this more implicit relationship to SCA has not yet entered the literature to the same extent as the aforementioned, more explicit goals related to SCA.

While SCA can have a positive marketable effect, practice does not primarily use it for better advertisement but to save costs. To 'record and utilize unexploited knowledge' and 'optimize existing products or services' evokes moderate to high ratings. Although the former goal is similarly, moderately represented in prior literature, the latter has been rarely addressed. While practitioners show that SCA adds value to their products, the literature does not specify for what purpose the methods can be used, even though use for product servitization or improved consultation services is possible.

Thus we find both similarities and differences in the goals and motivation to use SCA. Scientific literature, representing a research perspective, is not designed primarily to satisfy customers and instead seeks new insights or solutions to specific problems. But researchers should keep some practical motivations in mind when identifying research questions to pursue. In particular, the deviations that we identify suggest the need for more studies of the goal to optimize existing products or services, in line with studies that attempt to improve quality control of products or increase product quality, traceability, and reliability [99, 109, 110]. Beyond these applications, the underlying motivations imply some interesting research avenues. Practitioners regard SCA as highly relevant but struggle to implement it, and we need to determine if they are planning to implement SCA because it really provides benefits or simply because it seems "fancy" and can support marketing purposes. Another relevant avenue for research would involve the question of whether SCA actually fulfills the envisaged goal and satisfies practitioners' hopes about the benefits to be gained from the use of advanced analytics.

5.2. Analytics

Figure 17, which depicts the comparison between the MLR and practitioners' responses, reflects insights about the perceived relevance (responses that mentioned medium or high relevance) and the degree of implementation (responses that mentioned medium or high degrees), according to each analytics type and described application.

We again find a gap between the perceived relevance and degree of implementation of advanced analytics types. However, when we compare the survey findings with our MLR results, we also uncover a second gap: Especially WL focuses more on advanced predictive and prescriptive analytics types (cumulative 77%), not only when compared to the state of implementation (cumulative 46%) but also to the perceived relevance of practitioners (cumulative 63%). WL strives to develop and showcase emerging technologies and SCA applications, and thus, the focus on more mature analytics types is not surprising. In contrast, GL is slightly more aligned with the relevant topics identified by practitioners. With a cumulative 61% focusing on more advanced predictive and prescriptive analytics types, it is ahead of practitioners' implementation status and is aligned with the relevance they assign to these analytics types. This finding underscores our methodological justification for using an MLR. Due to its research-oriented, mostly cutting-edge focus, WL moves beyond the current concerns of practitioners,

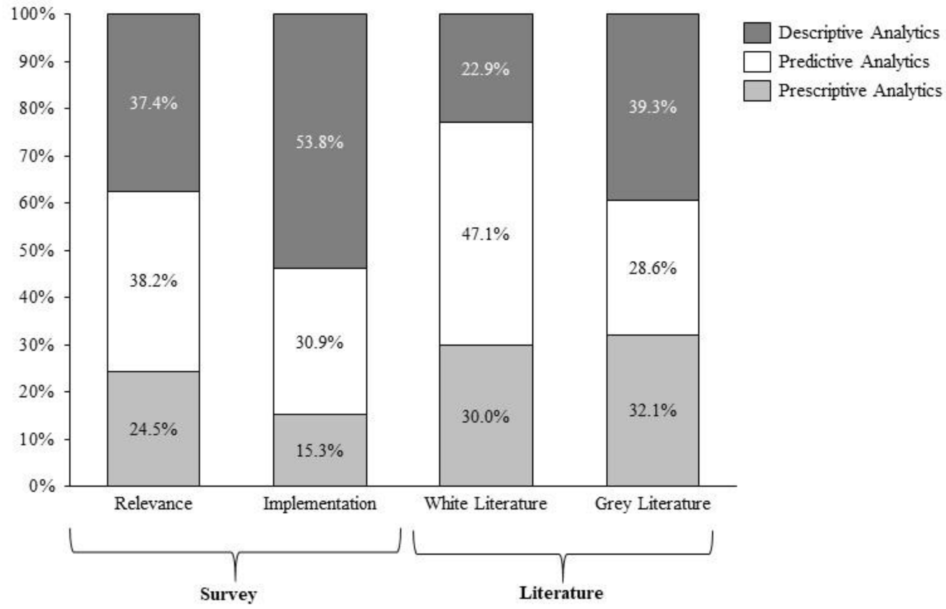

Figure 17: Consideration of analytics types by practitioners and literature

which is a natural and intended consequence of research publications. However, linking research results to how practitioners can apply them is an essential requirement for the successful usage of SCA in practice. GL positions itself in between WL and the practitioner's view because, while it needs to present innovations, it must still match its audience's more "down-to-earth" requirements. Nonetheless, similarly to WL, there is still an over-emphasis on prescriptive analytics in GL. The reason for that could be that value is only attained after a decision can be derived from SCA, which is achieved via the most mature analytics type. This, and the gap between perceived relevance and current implementation status, supports the impression that further work to aid practitioners in applying SCA is necessary.

In turn, we call for research that identifies the specific reasons for the gap in perceived relevance and degree of implementation, to clarify the barriers that hinder the implementation of more advanced SCA (e.g., data acquisition, missing skills and expertise) and find ways to overcome them. Conceivably, an SCA implementation guideline might be developed that acknowledges barriers, then carves a structured pathway that would enable supply chains to implement SCA. For descriptive analytics, practical implementation still lags behind its relevance, so we also recommend more research on descriptive analytics.

5.3. Data

Figure 18 shows the number of respondents that require a specific data type on the x-axis and the number of use cases from WL and GL on the x-axis, pertaining to the top ten data types.

Inventory and transport data are highly relevant for practitioners and prevalent in extant literature. This shows that inventory and transport data are the bread and butter of SCA. In the former case, literature is slightly more saturated than their actual practical relevance, but the opposite is true for transport data. Production planning data and production capacities are less practically significant, and their research coverage is moderate. Together with product master, phase-in/out, material flow disruption, and receipt data, they achieve a good fit, indicating medium to low practical relevance and moderate scientific coverage. In the bottom right quadrant, we find demand, quality, and goods receipt data, which have not been addressed in literature but represent strongly required data types according to the survey, where they represent the second to fourth most relevant types. For demand and quality data, this can be attributed to the high gap between relevant and available data which was investigated in subsection 4.2.3 which might also apply to the research situation. The evidence of underrepresentation of goods receipt data, however, which showed only a small gap between relevance and availability, requires additional investigation.

As we noted in subection 3.2.3, social media and website data are rare in GL, but the overall results also lead us to question to which extent they are really valuable in practical settings. Ease of accessibility might prompt researchers to use certain data types, but the value of implementing them in a practical use case might not justify the effort. Therefore, it would be helpful to research the value of data types for SCA. A definition of data costs, including data analysis efforts, information security, and privacy costs, could be a worthwhile effort. It then might support a comparison of value and cost and more transparent adoptions of some data in practice.

5.4. Application Area and Analytics

With a cross-analysis, we compare the relevance of different analytics types for each SCOR process, both for WL and GL and the expert survey. Unfortunately, a different interpretation of the SCOR processes became evident in our discussions with some survey participants. In the MLR, we classified papers to the second level of SCOR and distinguished different

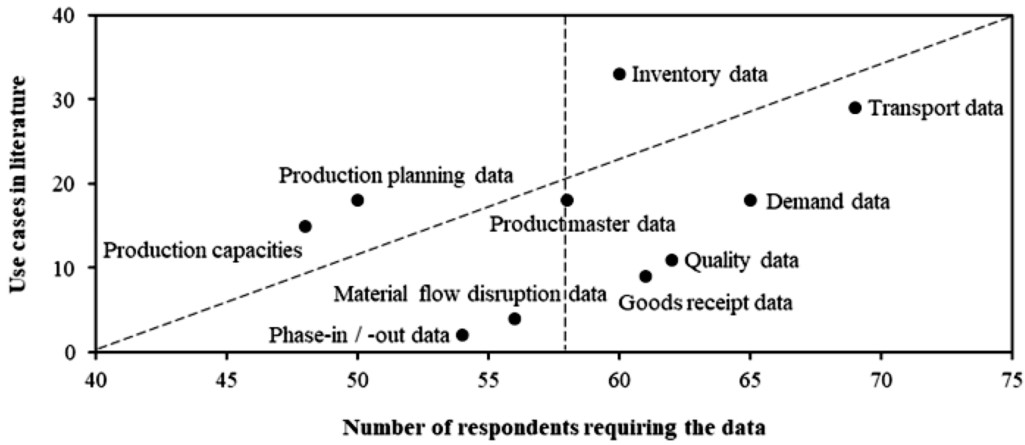

Figure 18: Consideration of data by practitioners and literature

planning processes, depending on what was planned (see Figure 19 for plan at the second level). This analysis reveals that the majority of publications discuss planning issues related to the whole supply chain (sP1 – Plan Supply Chain), such as demand forecasting.

The other plan sub-processes are meant to contain all planning activities related to each operational process (e.g., sP2 – Plan Source), but the survey participants appear to perceive such process-related planning activities as part of the respective operational process category (e.g., sS – Source). For example, the selection of suppliers should be in sP2, according to the definition provided by the American Production and Inventory Control Society [52]; it is a planning task related to the source process. But survey participants, possibly not aware of the detailed SCOR definitions, assign supplier selection to the source process, intuitively. Similar issues arise for plan make vs. make, plan deliver vs. deliver, and plan return vs. return.

Before we can compare the number of publications with the survey findings, we need to address this comprehension issue. In particular, we adjusted the category assignment of WL and GL, such that sP2 – Plan Source was presented as Source, sP3 – Plan Make changed to Make, sP4 – Plan Deliver became Deliver, and sP5 – Plan Return changed to Return. With these changes, the literature categorization aligns with the survey participants' interpretation of SCOR, so we can undertake a comparison of scientific and practical perceptions of relevancy. In scientific literature, relevance is represented by the number of sources that refer to an analytics type or SCOR process. In the survey, the respondents answered a direct question about their perceptions of relevance, on a scale from 0 (not relevant) to 5 (very relevant). We include answers that feature high or very high ratings, which form the basic population, so that they represent the 100% value for this analysis. Figure 20 plots the relevance of each SCOR process according to literature (WL and GL) and the survey, following the reinterpretation.

At a first glance, we note some seeming general consensus, but we also caution that survey participants did not assign values lower than 3.2 to the relevance of any SCOR process. Therefore, we identify a strong discrepancy in the return process. Survey participants only deem it slightly less relevant than make, source, or deliver; prior literature almost completely ignores it and provides few application cases. Especially noting the increasing interest in sustainability and efforts to find eco-friendly solutions, using SCA for return

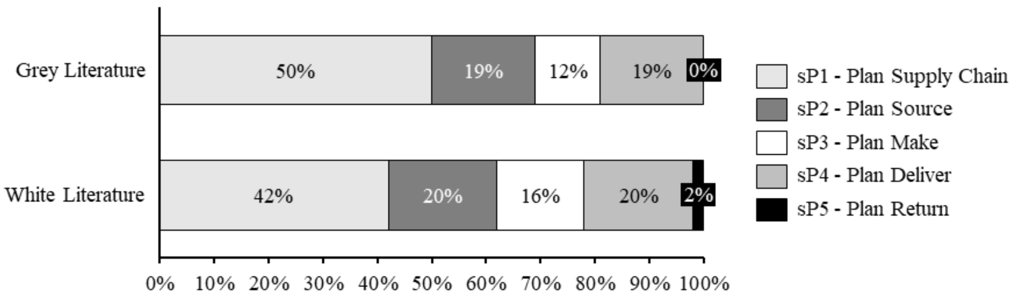

Figure 19: Division of SCOR Plan

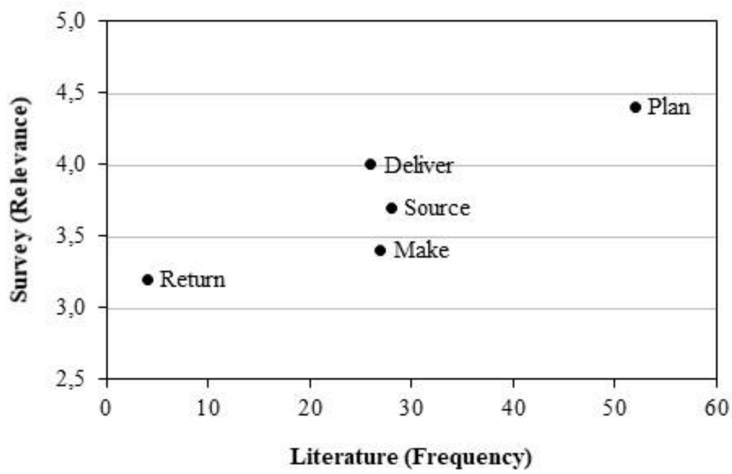

Figure 20: Perceived Relevance of SCOR processes

processes might be a promising field for research. In contrast, prior literature and the survey participants agree on the high relevance of plan processes. After reinterpreting the categories, as described, we identify a general agreement about the source, make, and deliver processes. They are regarded less relevant than plan, but the gap between the survey's relevance values (source 3.7, make 3.4, deliver 4.0) versus that for plan (4.4) is much smaller than the difference in publication frequency (source 28, make 27, deliver 26 vs. plan 52).

Considering SCOR processes and analytics types in relation, as in Figure 21, we find that the distribution of WL, GL, and survey responses regarding the relevance of each analytics type for each SCOR function, again, highlights the importance of planning, especially in WL. Notably, when WL applies descriptive analytics, it is mainly to solve planning tasks. Both GL and the survey apply descriptive analytics more evenly to all SCOR processes. These values have been normalized to sum to 100% for each bar, for ease of comparison, but the information about how often each analytics type is addressed is not reflected in this figure.

In general, the survey results assign about the same relevance to each analytics type for each SCOR process. Make and deliver are slightly underrepresented but considered more or less equally relevant for SCA. Industry respondents do not really differentiate among analytics types and their application areas, perhaps reflecting a lack of understanding of the strengths of each analytics type and the most suitable application cases. The survey confirms that practitioners regard analytics as (highly) relevant in general, but it also may imply that a wider range of application cases needs to be tested to provide clearer differentiation among analytics types. A Pearson correlation analysis could also confirm this notion, as the different analytics types had weak or no correlation between particular SCOR processes. In contrast, WL distinguishes the analytics types: Descriptive and predictive analytics are predominantly used for planning tasks, but prescriptive analytics use cases are more evenly distributed over all processes. A Phi correlation analysis could also support this: prescriptive analytics had almost no correlation to any SCOR process. Here, the coefficients only ranged from -0.084 (make) to 0.06 (plan). This finding might stem from the type of methods used for prescriptive analytics: Simulation and optimization are readily applicable to issues linked to operative SCOR processes, such as scheduling or routing. This rationale also might be why GL tends to focus on delivery, as the second-most addressed SCOR process. Considering the distribution of SCOR processes per analytics type, GL is situated somewhere between WL and the survey. Plan is represented more than other categories, but the focus is not as strong as in WL. This evidence supports the impression that GL examines topics closer to industry, which is not surprising. However, it must be remarked that the smaller sample size limits the expressiveness of findings and that is also the reason why no correlation analysis was conducted here.

The insights gained from comparing literature and practice interests regarding SCOR processes and a cross-analysis of SCOR processes and analytics types offer some further suggestions for research. First, very few investigations pertain to return processes. Practitioners think it is at least partly relevant for SCA (3.2 of 5 points), but WL includes only a few application cases. Noting ongoing discussions about critical topics such as sustainability, climate change, and resulting demands for all SCM functions, we posit that the relevance of a well-functioning, smooth return process will only increase, so research should focus on how SCA can contribute. Second, we need more research on operational processes in general, though it might be pertinent first to establish the reasons for a lack of interest thus far. If the reason is insufficient data to test application cases, stronger research-practice collaboration is required; if instead the few application

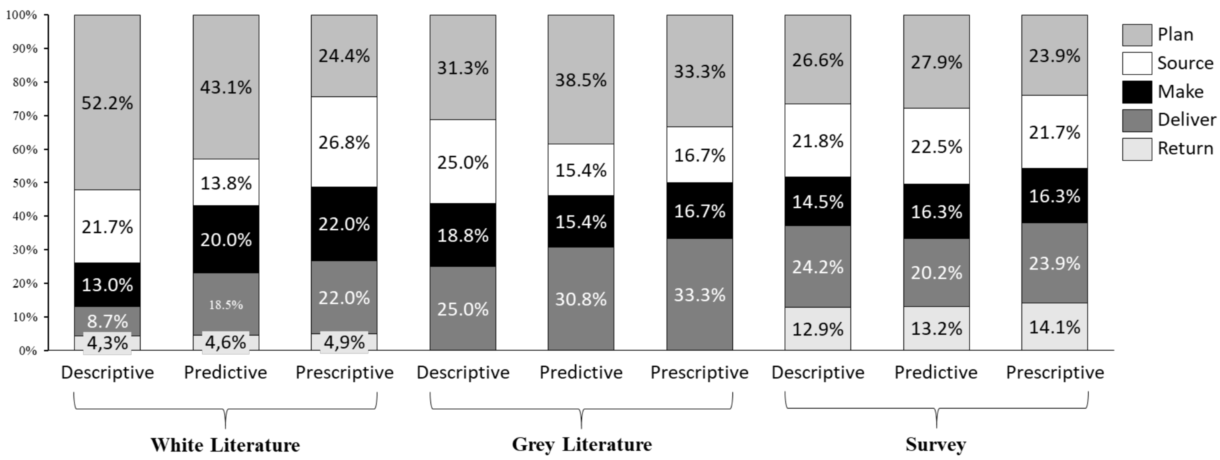

Figure 21: Relevance of SCOR processes and analytics types

cases reflect a lack of research opportunities, industry should promote the implementation of SCA for operative processes. Third, the cross-analysis indicates that practitioners do not distinguish among analytics types but instead assign each of them roughly the same relevance for each SCOR process.

Apparently, they have a rough understanding of what each type encompasses but lack detailed knowledge about the special strengths and application possibilities of each option. Instead of focusing on specific application cases, researchers might work to develop guidelines for when to apply which analytics type, which could help practice reap the benefits of SCA.

6. CONCLUSION

This article has sought to supply a holistic view of scientific-practice perspectives on SCA. Therefore, we conducted both an MLR and an expert survey to derive insights and also establish a comparison across sources. The structure for this analysis reflects the key components of SCA, as detailed in each section of this article: (1) the type of analytics employed, (2) the data, (3) the functional application area, and (4) the underlying goals and motivations for the application. In general, this study reveals a strong relevance of SCA, according to both literature-based investigations and a survey among practitioners. In addition to summarizing the key findings, we acknowledge some limitations and research opportunities in this section.

6.1. Summary

The main reason to use SCA is to improve decision-making. Both literature and practice focus on applying SCA to planning tasks, though practice prioritizes descriptive analytics, whereas literature emphasizes predictive analytics. A wide variety of data types are valuable for SCA. Although practitioners expect notable benefits from using SCA and regard it as relevant for various application scenarios, its implementation often does not live up to their expectations. Furthermore, the comparison reveals some critical divergences, such as the focus on different analytics types and the lack of differentiation across them, as well as diverse notions about potentially usable data for SCA and their application areas. The deviations offer some provisional insights for SCA research: Its core function continues to be pushing the cutting edge of the (technologically) possible, but creating links to practical value also is critical. In particular, research that assimilates practitioner-perceived value and scientific output is necessary. Such an intermediate step can be executed by different actors. For example, in GL, insights generated by research can converge to produce content that is easier for practitioners to leverage. The MLR provides additional insights into the field, which represents a valuable extension of traditional literature reviews.

Considering the link between research and practice is valuable for ensuring that we retain the value-adding potential of research results and enable their diffusion into practice. Carter [111] argues that the gaps between research and practice in SCM stem from knowledge transfer and knowledge production issues. To bridge the gaps and encourage the adoption of generated research findings, we specify some areas that should be the focus of efforts to transfer SCA research to maximize its impact.

6.2. Limitations

Although the WL review is extensive, the GL search relied on a narrow keyword string that explicitly noted “supply chain analytics.” A broader string might produce additional, relevant results, but when we attempted this step, the search engine returned almost exclusively WL. For GL, we also imposed restrictions regarding the file type to ensure the required quality level (e.g., first- or second-tier GL; [13]). But this methodological choice also might exclude some relevant first- or second-tier GL published in different file types, as well as third-tier literature that might contain promising results. Continued research could relax the restrictions on keywords, file type, and quality. Further, the relevance of the search engine is calculated by the Google PageRank algorithm, which can be biased and intransparent [112]. Still, we acquiesce to this drawback as we a) think that the benefits of including GL outweigh the downsides, b) perceive the Google search engine to be the best way to search for the targeted GL, and c) took precautions (i.e., no cookies, application of assessment criteria as proposed by Garousi et al. [13]). Lastly, the choice to not extend our search beyond 100 Google results is questionable. After all, relevant results could appear after these 100 results, especially considering the nature of the PageRank algorithm explained above. Nevertheless, it is strongly indicated that theoretical saturation has been achieved because the last 28 entries (more than 25% of all results) were irrelevant to our research.

In acquiring our experts we face different potential respondent biases. Among them is our focus on German speaking respondents as well as the BVL community in particular. Though we generated a representative sample of the BVL community, this community might not exactly match the industrial proportions of supply chain participants.

Some other limitations pertain to the categories used to analyze the findings. The goals and motivation for applying SCA were hard to compare across sources, because of their context specificity. The survey produced some interesting results, but the literature review rarely identifies goals explicitly and cites only a few particular motivations. Building robust hypotheses about goals and motivations of SCA on the basis of the literature review thus is unlikely. In addition, the category analytic types encompass three types (descriptive, predictive, and prescriptive), without detailing the specific method or algorithm used. For example, predictive analytics rely on both simple linear and complex neural network regressions, which require entirely different data, skills, and knowledge. An investigation of each SCA method may generate further insights.

In seeking to describe the gaps between science and practice, we devote less effort to explaining why these gaps exist. For example, we find that some analytics types are more frequently used and that some data are not available to practitioners, but we do not know why. Such conclusions would require further research. Additional expert or Delphi studies also could tackle concerns that arise because the SCOR definition was not well understood by the survey respondents. We rearranged the SCOR plan processes into process elements, to support comparisons with the survey, but the level of true comparability might be limited. Similarly, for other analyses, we measure the prevalence of certain analytics types or applications in prior literature, but the survey required respondents to rate the relevance and level of adoption of SCA, so the comparison assumes that prevalence is equivalent to relevance and adoption. In some cases, categories had a low literary prevalence, but no survey categories rated any of them below moderate, on average.

6.3. Further Research

Beyond efforts to address the limitations, this study reveals some interesting points for continued research. Regarding the analytics types, WL currently aligns with GL with regard to prescriptive analytics, but a stronger focus by scientific literature might foster expanded adoption in practice. For prescriptive and predictive analytics, it also might be beneficial to investigate the reasons for the low adoption rate in practice, despite their similar relevance to descriptive analytics (e.g., missing data, complex algorithms, missing use cases). Overall, the field lacks guidance regarding which analytics type practitioners should use and when.

For practitioners, we find it useful to consider the underlying drivers of their SCA initiatives to be able to choose a path for implementation compatible with the underlying incentive-structure (derived from the goals and motivations). We argue that further investigation of such a goal-implementation fit would enable practitioners to conduct their SCA projects with a higher degree of adoption success.

From a data standpoint, the comparison of GL with WL suggests the need for WL to provide more SCA research insights based on goods receipts, supply chain networks, inventory data, and production capacities. Other investigations should address why website data are so severely underrepresented in practice and what hinders their implementation. The survey results signal a big gap between the availability and accessibility of demand forecasts, quality, and material flow disruption data, so we need studies that can establish why these highly relevant data are often missing. A first analysis of the firm's interests to share data revealed that often unwillingness to share the information is the primary reason for the unavailability. Although demand, quality, and goods receipt data are relevant, they do not appear often in scientific literature, which suggests a pathway for research.

The MLR also reveals that four operational SCOR processes are poorly represented (source, make, deliver, return), such that the dominant focus is on the plan process. We argue that SCA for return processes in particular demands more attention, especially in relation to sustainability topics. To frame research efforts into other operative functions, we call for studies that identify the reasons for the current gaps. Finally, applying SCPM revealed a severe lack of SCA research for short-term production planning, whereas cross-functional planning is overrepresented in WL relative to GL. Thus, either some inhibitors are preventing adoption, or the relevance for practice is low, and this question should be addressed.

ACKNOWLEDGMENTS

The authors acknowledge crucial contributions by Lucas Stampe in the initial structuring of the article and the conceptualization of the literature review.

REFERENCES

[1] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: a survey," *Digital Communications and Networks*, vol. 4, no. 3, pp. 161–175, 2018.

[2] S. Tiwari, H. M. Wee, and Y. Daryanto, "Big data analytics in supply chain management between 2010 and 2016: Insights to industries," *Computers & Industrial Engineering*, vol. 115, pp. 319–330, 2016.

[3] T. Papadopoulos and K. Spanaki, "Data as an artefact in supply chain literature: A conceptual framing," in *OR60: The OR Society Annual Conference*, Lancaster, United Kingdom, pp. 84-93, 2018.

[4] J. P. U. Cadavid, S. Lamouri, and B. Grabot, "Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review," in *International Conference on Information Systems, Logistics and Supply Chain*, Lyon, France, 2018.

[5] V. Arya, P. Sharma, A. Singh, and P. de Silva, "An exploratory study on supply chain analytics applied to spare parts supply chain," *Benchmarking: An International Journal*, vol. 24, no. 6, pp. 1571–1580, 2017.

[6] J. Klein, R. Mello, A.-K. Cordes, B. Hellingrath, and R. Martins, "Big Data Analytics for Supply Chain Management: a review of empirical studies.," in *Proceedings of the 25th International Annual European Operations Management Association Conference*, Budapest, Hungary, 2018.

[7] M. Brinch, J. Stentoft, J. K. Jensen, and C. Rajkumar, "Practitioners understanding of big data and its applications in supply chain management," *The International Journal of Logistics Management*, vol. 29, no. 2, pp. 555–574, 2018.

[8] B. Chae, D. Olson, and C. Sheu, "The impact of supply chain analytics on operational performance: a resource-based view," *International Journal of Production Research*, vol. 52, no. 16, pp. 4695–4710, 2014.

[9] S. Chehbi-Gamoura, R. Derrouiche, D. Damand, and M. Barth, "Insights from big Data Analytics in supply chain management: an all-inclusive literature review using the SCOR model," *Production Planning & Control*, vol. 31, no. 5, pp. 355–382, 2020.

[10] D. Arunachalam, N. Kumar, and J. P. Kawalek, "Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice," *Transportation Research Part E: Logistics and Transportation Review*, vol. 114, C, pp. 416–436, 2018.

[11] T. Schoenherr and C. Speier-Pero, "Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential," *Journal of Business Logistics*, vol. 36, no. 1, pp. 120–132, 2015.

[12] G. Wang, A. Gunasekaran, E. W. Ngai, and T. Papadopoulos, "Big data analytics in logistics and supply chain management: Certain investigations for research and applications," *International Journal of Production Economics*, vol. 176, pp. 98–110, 2016.

[13] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Information and Software Technology*, vol. 106, pp. 101–121, 2019.

[14] B. K. Chae, D. Olson, and C. Sheu, "The impact of supply chain analytics on operational performance: A resource-based view," *International Journal of Production Research*, vol. 52, no. 16, pp. 4695–4710, 2014.

[15] B. K. Chae and D. L. Olson, "Business analytics for supply chain: A dynamic-capabilities framework," *International Journal of Information Technology and Decision Making*, vol. 12, no. 1, pp. 9–26, 2013.

[16] B. S. Sahay and J. Ranjan, "Real time business intelligence in supply chain analytics," *Information Management and Computer Security*, vol. 16, no. 1, pp. 28–48, 2008.

[17] H. Min, Y.-K. Lim, and J.-W. Park, "Supply chain analytics for enhancing the maritime security," *International Journal of Logistics Systems and Management*, vol. 28, no. 2, pp. 164–179, 2017.

[18] S. Zhu, C. Cegielski, J. Song, and K. B. Lee, "The impact of supply chain analytics on operational supply chain transparency: An information processing view," in *America's Conference on Information Systems Proceedings*, Boston, USA, 2017.

[19] S. Zhu, J. Song, B. T. Hazen, K. Lee, and C. Cegielski, "How supply chain analytics enables operational supply chain transparency: An organizational information processing theory perspective," *International Journal of Physical Distribution and Logistics Management*, vol. 48, no. 1, pp. 47–68, 2018.

[20] G. C. Souza, "Supply chain analytics," *Business Horizons*, vol. 57, no. 5, pp. 595–605, 2014.

[21] M. W. Barbosa, M. B. Ladeira, and A. de La Calle Vicente, "An analysis of international coauthorship networks in the supply chain analytics research area," *Scientometrics*, vol. 111, no. 3, pp. 1703–1731, 2017.

[22] B. Chae and D. L. Olson, "Business analytics for supply chain: A dynamic-capabilities framework," *International Journal of Information Technology & Decision Making*, vol. 12, no. 01, pp. 9–26, 2013.

[23] H. Min, Y.-K. Lim, and J.-W. Park, "Supply chain analytics for enhancing the maritime security," *International Journal of Logistics Systems and Management*, vol. 28, no. 2, pp. 164–179, 2017.

[24] S. Zhu, J. Song, B. Hazen, K. Lee, and C. Cegielski, "How supply chain analytics enables operational supply chain transparency," *International Journal of Physical Distribution & Logistics Management*, vol. 48, no. 1, pp. 47–68, 2017.

[25] S. Zhu, J. Song, B. T. Hazen, K. Lee, and C. Cegielski, "How supply chain analytics enables operational supply chain transparency," *International Journal of Physical Distribution & Logistics Management*, vol. 48, no. 1, pp. 47–68, 2018.

[26] A. Aryal, Y. Liao, P. Nattuthurai, and B. Li, "The emerging big data analytics and IoT in supply chain management: a systematic review," *Supply Chain Management: An International Journal*, vol. 25, no. 2, pp. 141–156, 2018.

[27] M. Awwad, P. Kulkarni, R. Bapna, and A. Marathe, "Big Data Analytics in Supply Chain: A Literature Review," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Washington DC, USA, pp. 418–425, 2018.

[28] A. Chakroun, A. E. Bouchti, and H. Abbar, "Logistics and Supply Chain Analytics: Benefits and Challenges," in *Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, London, UK, pp. 44–50, 2018.

[29] M. Moufaddal, A. Benghabrit, and I. Bouhaddou, "Big Data Analytics for Supply Chain Management," in *Lecture Notes in Networks and Systems, Innovations in Smart Cities and Applications*, M. Ben Ahmed and A. A. Boudhir, Eds., Cham: Springer International Publishing, pp. 976–986, 2018.

[30] T. Nguyen, L. Zhou, V. Spiegler, P. Ieromonachou, and Y. Lin, "Big data analytics in supply chain management: A state-of-the-art literature review," *Computers & Operations Research*, vol. 98, pp. 254–264, 2018.

[31] S. Fosso Wamba and S. Akter, "Big Data Analytics for Supply Chain Management: A Literature Review and Research Agenda," in *Lecture Notes in Business Information Processing, Enterprise and Organizational Modeling and Simulation*, J. Barjis, R. Pergl, and E. Babkin, Eds., Cham: Springer International Publishing, pp. 61–72, 2015a.

[32] S. Lodemann and W. Kersten, "Identifikation und Umsetzung der Potenziale von Data Analytics im Supply Chain Management," *Mittelstand-Digital Magazin Wissenschaft trifft Praxis*, no. 13, 22–29, 2020.

[33] K. Hoberg, M. Petersen, and J. Heinen, "Die Implikationen digitaler Technologien für die Supply Chain 4.0," in *Handbuch Industrie 4.0 und Digitale Transformation*, R. Obermaier, Ed., Wiesbaden: Springer Gabler, pp. 165–187, 2019.

[34] J. Loucks, T. Davenport, and D. Schatsky, "State of AI in the Enterprise: 2nd Edition," 2018. [Online]. Available: https://www2.deloitte.com/content/dam/insights/us/articles/4780_State-of-AI-in-the-enterprise/DI_ State-of-AI-in-the-enterprise-2nd-ed.pdf

[35] M. G. Alles, "Drivers of the use and facilitators and obstacles of the evolution of big data by the audit profession," *Accounting Horizons*, vol. 29, no. 2, pp. 439–449, 2015.

[36] E. M. Rogers, *Diffusion of innovations*. New York, NY, USA: Simon and Schuster, 2003.

[37] T. H. Kwon and R. W. Zmud, "Unifying the fragmented models of information systems implementation," in *Critical Issues in Information Systems Research*, R. J. Boland Jr and Hirschheim R. A., Eds., New York: Wiley, 1987.

[38] T. H. Davenport, "Analytics 3.0," *Harvard Business Review*, vol. 91, no. 12, pp. 64–72, 2013.

[39] T. T. Herden and S. Bunzel, "Archetypes of Supply Chain Analytics Initiatives – An Exploratory Study," *Logistics*, vol. 2, no. 10, pp. 1–20, 2018.

[40] N. Mishra and S. Silakari, "Predictive Analytics: A Survey, Trends, Applications, Oppurtunities & Challenges," *International Journal of Computer Science and Information Technologies*, vol. 3, no. 3, pp. 4434–4438, 2012.

[41] K. Lepenioti, A. Bousdekis, D. Apostolou, and G. Mentzas, "Prescriptive analytics: Literature review and research challenges," *International Journal of Information Management*, vol. 50, pp. 57–70, 2020.

[42] I. Varela Rozados and B. Tjahjono, "Big Data Analytics in Supply Chain Management: Trends and Related Research," in *6th International Conference on Operations and Supply Chain Management*, Bali, 2014.

[43] A. Gunasekaran et al., "Big data and predictive analytics for supply chain and organizational performance," *Journal of Business Research*, vol. 70, pp. 308–317, 2017.

[44] P. Anitha and M. M. Patil, "A Review on Data Analytics for Supply Chain Management: A Case study," *International Journal of Information Engineering and Electronic Business*, vol. 10, no. 5, pp. 30–39, 2018.

[45] G. C. Pflug, *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*. Boston, MA: Springer, 1996.

[46] M. K. Pusala, M. Amini Salehi, J. R. Katukuri, Y. Xie, and V. Raghavan, "Massive Data Analysis: Tasks, Tools, Applications, and Challenges," in *Big Data Analytics: Methods and Applications*, S. Pyne, B. P. Rao, and S. B. Rao, Eds., New Delhi: Springer India, pp. 11–40, 2016.

[47] V. Rajaraman, "Big data analytics," *Resonance*, vol. 21, no. 8, pp. 695–716, 2016.

[48] H. Hu, Y. Wen, T. Chua, and X. Li, "Toward Scalable Systems for Big Data Analytics: A Technology Tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.

[49] H. J. Watson, "Tutorial: Big Data Analytics: Concepts, Technologies, and Applications," *Communications of the Association for Information Systems*, vol. 34, 2014.

[50] H. Wenzel, D. Smit, and S. Sardesai, "A literature review on machine learning in supply chain management," in *Artificial Intelligence and Digital Transformation in Supply Chain Management: Innovative Approaches for Supply Chains*, Hamburg, Germany, pp. 413–441.

[51] U. Thonemann and F. Papier, "A1.4 Supply Chain Management," in vol. 3, *Handbuch Logistik*, D. Arnold, H. Isermann, A. Kuhn, H. Tempelmeier, and K. Furmans, Eds., Berlin: Springer, 2008.

[52] American Production and Inventory Control Society, *Supply chain operations reference model (Version 12.0)*. Chicago, IL, 2017.

[53] Stewart Gordon, "Supply-chain operations reference model (SCOR): the first cross-industry framework for integrated supply-chain management," *Logistics Information Management*, vol. 10, no. 2, pp. 62–67, 1997.

[54] F. B. Georgise, T. Wuest, and K.-D. Thoben, "SCOR model application in developing countries: challenges & requirements," *Production Planning & Control*, vol. 28, no. 1, pp. 17–32, 2017.

[55] B. Fleischmann, H. Meyr, and M. Wagner, "Advanced Planning," in *Supply Chain Management and Advanced Planning: Concepts, Models, Software, and Case Studies*, H. Stadtler, C. Kilger, and H. Meyr, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 71–95, 2015.

[56] K. M. Benzies, S. Premji, K. A. Hayden, and K. Serrett, "State-of-the-evidence reviews: advantages and challenges of including grey literature," *Worldviews on Evidence-Based Nursing*, vol. 3, no. 2, pp. 55–61, 2006.

[57] A. A. Chadegani et al., "A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases," *Asian Social Science*, vol. 9, no. 5, pp. 18–26, 2013.

[58] I. Hausladen and M. Schosser, "Towards a maturity model for big data analytics in airline network planning," *Journal of Air Transport Management*, vol. 82, no. 101721, pp. 1–18, 2020.

[59] R. Yanamandra, "A Framework of Supply Chain Strategies to achieve competitive advantage in Digital era," in *2019 International Conference on Digitization(ICD): Theme: Landscaping Artificial Intelligence*, Sharjah, United Arab Emirates, pp. 129–134, 2019.

[60] H. S. Kilic, A. E. Demirci, and D. Delen, "An integrated decision analysis methodology based on IF-DEMATEL and IF-ELECTRE for personnel selection," *Decision Support Systems*, vol. 137, no. 113360, pp. 1–19, 2020.

[61] P. Galetsi, K. Katsaliaki, and S. Kumar, "Big data analytics in health sector: Theoretical framework, techniques and prospects," *International Journal of Information Management*, vol. 50, pp. 206–216, 2020.

[62] S. Fosso Wamba and S. Akter, "Understanding supply chain analytics capabilities and agility for data-rich environments," *International Journal of Operations & Production Management*, vol. 39, 6/7/8, pp. 887–912, 2019b.

[63] B. Chae, C. Yang, D. Olson, and C. Sheu, "The impact of advanced analytics and data accuracy on operational performance: A contingent resource based theory (RBT) perspective," *Decision Support Systems*, vol. 59, pp. 119–126, 2014.

[64] A. M. T. Thomé, L. F. Scavarda, and A. J. Scavarda, "Conducting systematic literature review in operations management," P*roduction Planning & Contro*l, vol. 27, no. 5, pp. 408–420, 2016.

[65] K. E. White, *Publications Output: U.S. Trends and International Comparisons*. [Online]. Available: https://ncses.nsf.gov/pubs/nsb20206/publication-output-by-region-country-or-economy

[66] G. Walter, F. Cordes, J. Rodriguez, J. Lowe, and N. Panday, "Turning Visibility into Value in Digital Supply Chains," BCG, 2018. [Online]. Available: https://image-src.bcg.com/Images/BCG-Turning-Visibility-into-Value-in-Digital-Supply-Chains-Jan-2018_tcm9-181967.pdf

[67] P. Tawde and S. Jaswal, "Optimization of SCM process using evolutionary algorithm on SVR," in *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)*, Chirala, Andhra Pradesh, India, pp. 103–106, 2017.

[68] N. Q. Viet, B. Behdani, J. Bloemhof, and K. Hoberg, "Value of data in multi-level supply chain decisions: a case study in the Dutch floriculture sector," *International Journal of Production Research*, vol. 59, no. 5, pp. 1–18, 2020.

[69] F. H. Gjendem and Deep Aakash, "Data-Driven and Analytics-Powered: Digital supply chain is the new reality," Accenture, 2016. [Online]. Available: https://www.accenture.com/_acnmedia/PDF-26/Accenture-Digital-Supply-Chain-Operations-Analytics-POV.pdf

[70] C.-H. Wang and Y. Yun, "Demand planning and sales forecasting for motherboard manufacturers considering dynamic interactions of computer products," *Computers & Industrial Engineering*, vol. 149, no. 106788, pp. 1–8, 2020.

[71] L. C. Boldt et al., “Forecasting Nike’s sales using Facebook data,” in *2016 IEEE International Conference on Big Data (Big Data)*, Washington DC,USA, pp. 2447–2456, 2016.

[72] M. A. Villegas, D. J. Pedregal, and J. R. Trapero, “A support vector machine for model selection in demand forecasting applications,” *Computers & Industrial Engineering*, vol. 121, pp. 1–7, 2018.

[73] A. Grackin, L. Rosencrance, and D. Turbide, “Grow Profits With Supply Chain Analytics,” TechTarget, 2016. [Online]. Available: http://docs.media.bitpipe.com/io_13x/io_131640/item_1340667/HB_Grow%20Profits%20With%20Supply%20Chain%20Analytics_final.pdf

[74] S. Rowe and M. Pournader, “Supply Chain Big Data Series Part 1: How big data is shaping the supply chains of tomorrow,” KPMG, 2017. [Online]. Available: https://assets.kpmg/content/dam/kpmg/au/pdf/2017/big-data-analytics-supply-chain-performance.pdf

[75] S. Hosseini and A. A. Khaled, “A hybrid ensemble and AHP approach for resilient supplier selection,” *Journal of Intelligent Manufacturing*, vol. 30, no. 1, pp. 207–228, 2019.

[76] A. Kim, J. Obregon, and J.-Y. Jung, “PRANAS: A Process Analytics System Based on Process Warehouse and Cube for Supply Chain Management,” *Applied Sciences*, vol. 10, no. 3521, pp. 1–13, 2020.

[77] P.-J. Wu, M.-C. Chen, and C.-K. Tsau, “The data-driven analytics for investigating cargo loss in logistics systems,” *International Journal of Physical Distribution & Logistics Management*, vol. 47, no. 1, pp. 68–83, 2017.

[78] L. Li, S. Ma, X. Han, C. Zheng, and Di Wang, “Data-driven online service supply chain: a demand-side and supply-side perspective,” *Journal of Enterprise Information Management*, vol. 34, no. 1, pp. 1–17, 2020.

[79] I. M. Cavalcante, E. M. Frazzon, F. A. Forcellini, and D. Ivanov, “A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing,” *International Journal of Information Management*, vol. 49, pp. 86–97, 2019.

[80] K. M. Sabu and T. M. Kumar, “Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala,” *Procedia Computer Science*, vol. 171, pp. 699–708, 2020.

[81] R. Sathyan, P. Parthiban, R. Dhanalakshmi, and A. Minz, “A combined big data analytics and Fuzzy DEMATEL technique to improve the responsiveness of automotive supply chains,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 7, pp. 7949–7963, 2020.

[82] M. I. Ali, P. Patel, and J. G. Breslin, “Middleware for Real-Time Event Detection and Predictive Analytics in Smart Manufacturing,” in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, Santorini Island, Greece, pp. 370–376, 2019.

[83] S. Punia, S. P. Singh, and J. K. Madaan, “A cross-temporal hierarchical framework and deep learning for supply chain forecasting,” *Computers & Industrial Engineering*, vol. 149, no. 106796, pp. 1–11, 2020.

[84] T. Sandiford, “Supply Chain Analytics,” Capgemini, 2013. [Online]. Available: https://www.capgemini.com/wp-content/uploads/2017/07/supply_chain_analytics_0.pdf

[85] E. Malinowski, M. H. Karwan, J. M. Pinto, and L. Sun, “A mixed-integer programming strategy for liquid helium global supply chain planning,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 110, pp. 168–188, 2018.

[86] H. Gholizadeh, H. Fazlollahtabar, and M. Khalilzadeh, “A robust fuzzy stochastic programming for sustainable procurement and logistics under hybrid uncertainty using big data,” *Journal of Cleaner Production*, vol. 258, no. 120640, pp. 1–24, 2020.

[87] S. Mishra and S. P. Singh, “Distribution network model using big data in an international environment,” *The Science of the total environment*, vol. 707, no. 135549, pp. 1–11, 2020.

[88] G. Sankaran, M. Knahl, G. Siestrup, and I. Vasileiou, “Value of Smart Data for Supply Chain Decisions in a Data Rich, Uncertain World,” in *Proceedings of the 5th Collaborative European Research Conference (CERC 2019)*, 2019.

[89] V. Jain, S. Kumar, A. Kundu, and C. Chandra, “Linking procurement operations to power-influence in a supply chain,” *Journal of Manufacturing Systems*, vol. 41, pp. 201–210, 2016.

[90] B. Shen, T.-M. Choi, and H.-L. Chan, “Selling green first or not? A Bayesian analysis with service levels and environmental impact considerations in the Big Data Era,” *Technological Forecasting and Social Change*, vol. 144, pp. 412–420, 2019.

[91] C. K. H. Lee, "A GA-based optimisation model for big data analytics supporting anticipatory shipping in Retail 4.0," *International Journal of Production Research*, vol. 55, no. 2, pp. 593–605, 2017.

[92] Q. Cao, B. Bouqata, P. D. Mackenzie, D. Messier, and J. J. Salvo, "A grid-based clustering method for mining frequent trips from large-scale, event-based telematics datasets," in *2009 IEEE International Conference on Systems, Man and Cybernetics*, San Antonio, TX, USA, pp. 2996–3001, 2009.

[93] IBM, "IBM Cognos Supply Chain Analytics," IBM, 2009. [Online]. Available: https://silo.tips/download/white-paper-february-ibm-cognos-supply-chain-analytics

[94] WMG Business, "Supply chain analytics helps Tetley thrive despite pandemic," University of Warwick, 2020. [Online]. Available: https://warwick.ac.uk/fac/sci/wmg/research/transformation/supply_chain/wmg_tetley_success_story_supplyvue.pdf

[95] R. Iftikhar and M. S. Khan, "Social Media Big Data Analytics for Demand Forecasting: Development and Case Implementation of an Innovative Framework," *Journal of Global Information Management*, vol. 28, no. 1, pp. 103–120, 2020.

[96] R. Wen and W. Yan, "Supply–Demand Prediction for Agile Manufacturing with Deep Neural Network," *Smart and Sustainable Manufacturing Systems*, vol. 3, no. 2, pp. 95–105, 2019.

[97] S. Papagiannidis, E. W. See-To, D. G. Assimakopoulos, and Y. Yang, "Identifying industrial clusters with a novel big-data methodology: Are SIC codes (not) fit for purpose in the Internet age?," *Computers & Operations Research*, vol. 98, pp. 355–366, 2018.

[98] N. K. Dev, R. Shankar, R. Gupta, and J. Dong, "Multi-criteria evaluation of real-time key performance indicators of supply chain with consideration of big data architecture," *Computers & Industrial Engineering*, vol. 128, pp. 1076–1087, 2019.

[99] M. Mancini, A. Mircoli, D. Potena, C. Diamantini, D. Duca, and G. Toscano, "Prediction of pellet quality through machine learning techniques and near-infrared spectroscopy," *Computers & Industrial Engineering*, vol. 147, no. 106566, 2020.

[100] E. Isikli and M. Ketenci, "Using Consumer Reviews for Demand Planning: Case of Configurable Products," in *Advances in Intelligent Systems and Computing, Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making*, C. Kahraman, S. Cebi, S. Cevik Onar, B. Oztaysi, A. C. Tolga, and I. U. Sari, Eds., Switzerland: Springer Nature, pp. 354–361, 2020.

[101] A. Singh, N. Shukla, and N. Mishra, "Social media data analytics to improve supply chain management in food industries," *Transportation Research Part E: Logistics and Transportation Review*, vol. 114, pp. 398–415, 2018.

[102] R. Perumalsamy and J. Natarajan, "Predictive analytics using Genetic Algorithm for efficient supply chain inventory optimization," in *Second International conference on Computing, Communication and Networking Technologies*, Karur, India, pp. 1–8, 2010.

[103] M. Åkerman *et al.*, "Challenges Building a Data Value Chain to Enable Data-Driven Decisions: A Predictive Maintenance Case in 5G-Enabled Manufacturing," *Procedia Manufacturing*, vol. 17, pp. 411–418, 2018.

[104] S. van der Spoel, C. Amrit, and J. van Hillegersberg, "Predictive analytics for truck arrival time estimation: a field study at a European distribution centre," *International Journal of Production Research*, vol. 55, no. 17, pp. 5062–5078, 2017.

[105] K. Leng, L. Jing, I.-C. Lin, S.-H. Chang, and A. Lam, "Research on mining collaborative behaviour patterns of dynamic supply chain network from the perspective of big data," *Neural Computing and Applications*, vol. 31, S1, pp. 113–121, 2019.

[106] Genpact, "Driving supply chain excellence through Data-to-Action Analytics," Genpact, 2015. [Online]. Available: http://www.genpact.com/downloadable-content/insight/driving-supply-chain-excellence-through-data-to-action-analytics.pdf

[107] BVL, *Jahresbericht der Regionalgruppen 2017*. Bremen, 2017.

[108] BVL, *Jahresbericht der Regionalgruppen 2020*. Bremen, 2020.

[109] Y. Oh, M. Busogi, K. Ransikarbum, D. Shin, D. Kwon, and N. Kim, "Real-time quality monitoring and control system using an integrated cost effective support vector machine," *Journal of Mechanical Science and Technology*, vol. 33, no. 12, pp. 6009–6020, 2019.

[110] M. A. Benatia, D. Baudry, and A. Louis, "Detecting counterfeit products by means of frequent pattern mining," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2020.

[111] C. R. Carter, "Knowledge production and knowledge transfer: closing the research-practice gap," *Journal of Supply Chain Management*, vol. 44, no. 2, pp. 78–83, 2008.

[112] A. N. Langville and C. D. Meyer, *Google's PageRank and beyond: The science of search engine rankings*. Princeton, NJ: Princeton Univ. Press, 2006.

APPENDIX

A. REASONING FOR INCLUDING GREY LITERATURE

Table 6. Decision aid for including GL based on [36]

	Question	Answer
1	Is the subject "complex" and not solvable by considering only the formal literature?	Yes
2	Is there a lack of volume or quality of evidence, or a lack of consensus of outcome measurement in the formal literature?	Yes
3	Is the contextual information important to the subject under study?	No
4	Is it the goal to validate or corroborate scientific outcomes with practical experiences?	Yes
5	Is it the goal to challenge assumptions or falsify results from practice using academic research or vice versa?	Yes
6	Would a synthesis of insights and evidence from the industrial and academic community be useful to one or even both communities?	Yes
7	Is there a large volume of practitioner sources indicating high practitioner interest in a topic?	Yes

B. SURVEY QUESTIONS

Question underlying Figure 11:	“Which economic sector is your company part of?” A) Manufacturing B) Logistics Services C) Trade D) Other
Question underlying Figure 11:	“How many employees did your company have in the last business year?” A) Less than 10 B) 10-49 C) 50-249 D) 250-499 E) 500-2.999 F) 3.000-9.999 G) 10.000-19.999 H) More than 20.000
Question underlying Figure 12:	“What goals do you pursue with the use of Advanced Data Analytics?” A) Optimize internal processes B) Optimize decisions C) Optimize existing products/services D) Develop new products and services E) Acquire new customers F) Capture and use previously unused knowledge
Question underlying Figure 13:	“What is your motivation for initiating an Advanced Data Analytics project?” A) Use available data to add value B) Solve a specific issue C) Apply new technological possibilities to communicate this internally D) Apply new technological possibilities to communicate this externally E) Increase transparency/understanding for an area F) Draw level with competitors who already use data analysis G) Our partners demand the commitment
Question underlying Figure 14, 15, 16, 18, 19 & 22:	“Please evaluate how relevant the following technology concepts for data analysis / artificial intelligence are for SCM and logistics in your company and what the implementation status is in your company.” Relevance to the company… A) unknown concept B) very low C) low D) medium E) high F) vey high Implementation status in the company… A) not planned B) planned >5 years C) planned <5 years D) already implemented to a minor extent E) already implemented partially F) already implemented to a broad extent
Question underlying Figure 17 & 20:	“Please choose which data your company needs from your supply chain partners and which data your company would be willing to share with your supply chain partners.” Data we need from our partners: A) Yes, they are already available B) Yes, but they are not available C) No Data we would be willing to share with our partners: A) Yes, we are already sharing these B) Yes, but we do not share them yet C) No D) Does not apply to our company
Question underlying Figure 18, 22 & 23:	“How do you assess the potential for Advanced Data Analytics in the following areas of your company?” A) Very low B) low C) medium D) high E) very high F) don’t know / not specified