Big data analytics and demand forecasting in supply chains: a conceptual analysis

Abstract

Purpose – Demand forecasting is a challenging task that could benefit from additional relevant data and processes. This paper examines how big data analytics enhances forecasts' accuracy.

Design/methodology/approach – A conceptual structure based on the design-science paradigm is applied to create categories for big data analytics. Existing theories from the scientific literature are synthesized with industry knowledge through experience and intuition. Accordingly, a reference frame is developed using three steps: (i) description of conceptual elements of the frame utilizing justificatory knowledge, (ii) specification of principles of the theoretical frame to explain the interplay between elements, and (iii) creation of a matching frame by conducting investigations within the retail industry.

Findings – The developed framework could serve as a first guide for meaningful big data analytics initiatives in the supply chain. The paper illustrates that integration of different data sources in demand forecasting is feasible but requires data scientists to perform the job, an appropriate technological foundation, and technology investments.

Originality/value – So far, no scientific work has analyzed the relation of forecasting methods to big data analytics; previous works have described technologies, types of analytics, and forecasting methods separately. This paper, in contrast, combines insights and provides advice on how enterprises can employ big data analytics in their operational, tactical, or strategic demand plans.

Keywords – Big data; Analytics; Forecasting methods; Demand influencing factor; Retail supply chains

Paper type – Conceptual Paper

1 Introduction

Retailers know a lot about end consumers, perhaps more than we know ourselves. In 2012, for example, US retailer Target sent coupons for baby clothes to a high-school girl. Her father was not amused by this type of advertisement, decided to visit a local Target store and requested to see the person in charge. The manager apologized for the inappropriate advertisement then a few days later over the phone. However, the father, seemingly embarrassed, revealed during this phone conversation that he had talked to his daughter: "It turns out there's been some activities in my house I haven't been completely aware of. She's due in August" (Duhigg, 2012).

As market expectations, competition, and volatility are rising, retailers are exploring data analytics to address new challenges and opportunities. Data analytics techniques not only provide single companies with greater accuracy, clarity, and insight but also lead to more contextual "intelligence" shared across all supply chains regardless of industry or sector.

1.1 Initial situation

The example of US retailer Target illustrates that retailers possess a wealth of information about the market and their customers. This data is a valuable resource as it enables companies to strengthen their intermediary position between manufacturers and consumers (Zhan *et al.*, 2016). If retailers employ analytics, data becomes a factor that creates added value. How information is used to generate a competitive advantage is therefore strongly dependent on the method, scope, and purpose of data processing (Hazen *et al.*, 2014). For instance, Amazon, a leading online retailer, successfully patented what it calls "anticipatory shipping," a method in which deliveries are initiated prior to the actual customer order placement in order to cut delivery time. Amazon will box and ship products to areas where it anticipates orders will be placed. The packages may remain at a local hub or be sent to another hub near an active customer that is about to order (Bensinger, 2014). Another example is Otto, a German online and catalog retailer. The company underwent a cultural change and now employs data-driven decision-making with big data and business intelligence (BI). It uses predictive analytics to dramatically reduce the return rates of fashion items (Clarke, 2013).

Enterprises use different methods to overcome the challenges of forecasting product demand (Hazen *et al.*, 2016). Some methods are qualitative (e.g., market research or expert estimations), while others are quantitative (e.g., time-series forecasts). Problems regarding decision support are based on the use of specific methods and, in turn, are dependent upon specific input data.

In order to provide practical insights of big data in this study, an approach to its integration in demand forecasting is necessary. Unfortunately, the current literature is either too vague, driven by the wishful thinking of practitioners and consultants, or too technical and specific. Many scientific contributions focus on quantitative reasoning or mathematical approaches (Chen *et al*., 2015; Silva and Reilly, 2014), which require expertise and prior domain knowledge. There is still a lack of comprehensive frameworks enabling convenient access to big data in academia and in practice (Hazen *et al*., 2016; Wang et al., 2016; Akter and Wamba, 2016, Wamba *et al*., 2015).

1.2 Methodology

The paper at hand addresses the following research question: How can big data analytics improve demand forecasting? In the following, Whetten's (1989) three key research questions are elaborated upon to stress the relevance of this research and to describe the methodology chosen in this paper.

Why? First, there is no systematic overview of the available big data analytics (BDA) techniques and their relation to forecasting methods. Moreover, there is no generic conceptual framework to illustrate how data and information relate to the decision problem and situation from a systems point of view (Tan *et al*., 2015). Second, the literature seldom depicts the process of generating insight from various types of data. Thus, until now, there has been no comprehensive description of the mechanisms and applications that relate to the supply chain. Third, there is no match between types of big data analytics and forecasting methods although BDA uses extensive data to support decision-making in an accurate and generic manner. It is unclear if BDA is able to substitute or complement existing forecasting methods. Fourth, demand forecasts do not only affect short-term decisions; they also support medium- and long-term decisions, such as routing or site location problems. Thus, combination of BDA and certain forecasting methods makes it necessary to consider different time horizons.

To investigate the potential for big data to be applied to the supply chain (Wang *et al*., 2016), we focus on the retail industry. As the intersection between manufacturers and (end) consumers, the downstream part of the supply chain is traditionally characterized by a level of uncertainty and the direct effects of demand planning on overall turnover (Richard *et al*., 2012). Retailers need to make decisions such as whether to introduce new products (Zhan *et al*., 2016) and whether to open new shop locations (Fernandes *et al*., 2015). The benefits of more accurate consumer demand predictions in relation to these decisions are apparent. For example, enhanced predictions would improve replenishment forecasts, directly affecting retailers' profitability (Chase, 2014). Inventory is an important driver of cost for retailers as it ties up capital and uses storage space. However, limited product availability can have negative effects, such as loss of revenue, or diminished customer loyalty. "Out-of-stock" research (i.e., Hanson *et al*., 2015; Che *et al*., 2012; Fernie and Grant, 2008) emphasizes that product availability is a critical factor affecting customer satisfaction. An appropriate forecasting solution would generate the exact daily demand to ensure that there is the necessary amount of stored goods without excess. Predictable demand helps to organize orders, determine the shop's product assortment and placement, and manage order shipping, scheduling, and production. Consequently, the efficiency of the whole supply chain would be improved by more accurate forecasts.

What? The paper aims to provide a conceptual overview of big data analytics by introducing a systematic framework that illustrates how enterprises could approach specific technologies and their potential applications in relation to existing forecasting methods. The objective of the underlying study is to analyze, conduct, and determine the practical value of BDA in demand forecasting for the retail industry. While Blackburn *et al*. (2015) point out that traditional methods of demand forecasting are challenged by the ever-increasing volatility, uncertainty, complexity, and ambiguity (VUCA) of events in the supply chain, state-of-the-art techniques increase companies' competitive advantage (Lowson, 2001).

By answering the research question, the paper at hand makes the following contributions to related literature. First, we examine the relationships between BDA and forecasting. We assume that specific analytic types fit "better" with certain types of forecasting methods and that only alignment between the input, scope, and method ensures appropriate analysis of data to support decision-making for demand planning and forecasting. Second, the paper investigates the practical value of BDA. Certain methods not only support conventional forecasting but also replace traditional methods. Existing methods are compared to aggregate insights through inductive reasoning. Third, systematization of BDA and forecasting methods is conducted and it is emphasized that one "size" does not fit all; different types of analytics are appropriate for different forecasting situations and time horizons. By discussing the strategic, tactical, and operational aspects of BDA in retail supply chains, we determine the specific value of BDA techniques.

How? The topic of this paper concerns information systems research for operations management applications. In order to answer the research questions, conceptual and practical approaches are chosen. Along with additional study in the form of basic desk research and a literature review, we use a research method based on Gregor and Jones' (2007) design science theory. Building upon elements of design science, it is first necessary to define the term "big data analytics" before developing a reference frame for BDA in relation to demand forecasting. As the potential of big data remains unclear, this paper aims to extend knowledge by creating new and innovative "artifacts" that could lead to new management models, frameworks, and theories (Hevner *et al.*, 2004). Whereas artifacts facilitate productive application and management of big data technologies, the theoretical approach uses an iterative process to create a coherent reference frame for BDA and demand forecasting. Beginning by gathering domain knowledge, we then follow Meredith (1993) and define the practices, technical capabilities, and possible outcomes of BDA techniques within the supply chain. This approach allows us to systematically fill the identified knowledge gaps with conceptual descriptions, taxonomies, and typologies (Simon, 1996). After elaborating upon the preliminary insights, we repeat the cycle until the findings possess a certain construct validity (Hevner *et al.*, 2004). As a result, the conceptual reference frame is established through a three-stage process involving (i) explanation of elements using justificatory knowledge (in our case BDA techniques and demand forecasting methods), (ii) definitions of basic principles to outline the interactions between these elements within the theoretical frame, and (iii) achievement of a match of the elements and presentation of exemplary approaches from a practitioner's perspective (in our case the retail industry).

An essential assumption of our research is that the various forecasting methods we use are not likely to disappear in the developing paradigm of BDA but will certainly be adapted and improved. Nevertheless, our study accounts for a possible shift from one preferable method to another. The underlying idea of this paper is not to provide a complete overview of all current BDA techniques, but to present an orientation of how BDA implementations could improve demand forecasting.

The remainder of this paper is structured as follows. Section 2 provides the theoretical background, highlighting relevant conceptual elements like demand influencing factors, forecasting methods and BDA techniques. Section 3 introduces the frame to demonstrate the theoretical interplay between BDA and demand forecasting. In section 4, the framework is applied to retail supply chains. The paper closes with a conclusion and recommendations in section 5.

2 Background – the conceptual elements

Here, we provide an overview of the current scientific knowledge on demand influencing factors. Then, existing forecasting methods will be briefly discussed. Finally, the term "big data" will be elaborated upon. Since there is no general match between these complementary techniques, the paper focuses on details about BDA to close the knowledge gap.

2.1 Demand influencing factors

Supply chains can be seen as the underlying link between processes and parties to fulfill a customer's request (Horvath, 2001). These interrelated activities involve complex decision-making under uncertainty and risk. Supply chains are confronted with the bullwhip effect, in which inaccurate forecast decisions and information asymmetry yield inefficiencies. As a result of insufficient information, order quantities are often shifted from the retailer further up the supply chain (Barlas and Gunduz, 2011). Because of global operations, shortened product cycles, and more volatile business environments, the management of uncertainties is one priority for the market-driven demand chain (Christopher and Peck, 2003). Demand forecasting is used to anticipate future sales. Beside production planning, inventory management, market entry strategies, and analysis of customer behavior, different demand forecasting methods allow for prediction of probable scenarios based on historical data and prevailing trends. A precise demand forecast provides an accurate picture of future demand and helps to avoid overproduction and excessive overstock (Chambers *et al.*, 1971).

Although forecasting methods allow quantification of future demand, supply chain uncertainty refers to situations in which the decision maker acts under a set of different forces. These forces can be further specified to understand the relationships and effects that play a strong role in planning, improving efficiency, or obtaining accurate forecasts (Sinha *et al.*, 2016; Hämäläinen, 2011). Figure 1 illustrates an overview of the factors influencing demand (Croxton *et al.*, 2002). These factors are embedded into a customer-oriented supply chain. Uncertainties in demand forecasting depend on the objective of planning. The state of a product, determined by the "four Ps" in the marketing field, has an important relation to the underlying goal of fulfilling a customer request (Yazdanparast *et al.*, 2010). Important parameters for demand forecasting include the *product* (e.g., for identification of future demand), *placement* (e.g., for determining the modes of transport), *pricing* (e.g., for implementation of a new product), and *promotion* (e.g., for determining the target market) (Deshmukh and Mohan, 2016). In addition, the supply chain is associated with constraints that directly influence the forecasting process. The supply side as well as the demand side not only affect how information is processed but also define the sufficiency and quality of information (Galaso *et al.*, 2009; Wong *et al.*, 2011). Finally, the environment must be considered in demand forecasts as the role of competitors and external factors, such as weather conditions or special events, affect the efficacy of control actions (Flynn *et al.*, 2016).

Consideration of all these elements is needed to determine the appropriate demand forecasting method (Lundholm, 2010).

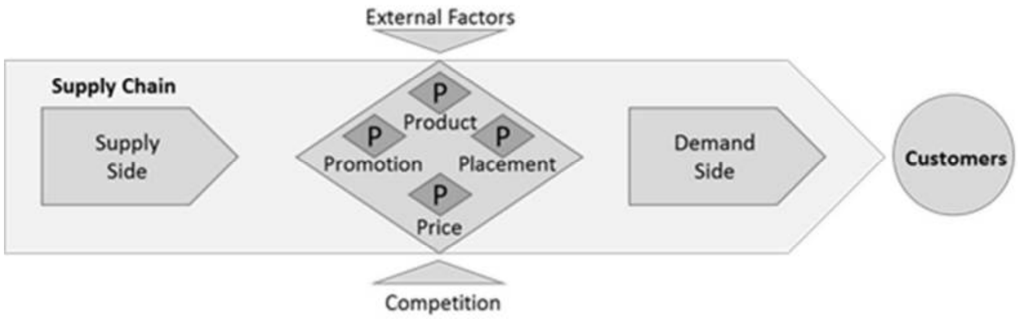

Figure 1: Schematic overview of factors influencing demand

2.2 Forecasting methods

Today, a wide array of forecasting methods are available, from which decision makers and analysts can select the most appropriate model. The literature roughly classifies forecasting methods into two main categories: qualitative and quantitative (Chambers *et al*., 1971; Armstrong, 2001; Makridakis *et al*., 2008). Qualitative forecasts, on the one hand, utilize qualitative data, such as expert opinions or knowledge of special events. They are primarily applied if the available data is inadequate for a quantitative analysis or if qualitative information is likely to increase the accuracy of forecasts (Hazen *et al*., 2016; Armstrong and Green, 2014). Quantitative methods, on the other hand, focus on patterns and pattern changes in historic data (time-series) or on specific relationships between system elements, such as causal models (Chambers *et al*., 1971).

Beyond the qualitative/quantitative classification scheme, there are five forecasting methods: (1) grassroots, (2) market research, (3) expert estimations, (4) time-series, and (5) causal models. This classification was developed by Armstrong (2001), a forecasting expert, and founder of the *Journal of Forecasting* and *International Journal of Forecasting*. It is based on the works of Martino (1977) and Bright (1978).

To answer the research question, it is important to determine which information is required for forecasting and what challenges forecasters face. Challenges can typically arise from a lack of data or skills, excessive costs, imperfect methods, or time restrictions (Chambers *et al*., 1971; Armstrong, 2001).

Several methods haven been developed, analyzed, and discussed in the general context of operations research and the specific context of retail management. Although all approaches have their individual strengths, they also have weaknesses. However, if and how these approaches are able to overcome their challenges through the utilization of BDA and discussion of the factors influencing retail business have not been debated. An overview of the forecasting methods is presented in Table 1, which also includes information about their challenges, which hint at the potential applications of BDA.

	Forecasting method	Required information	Challenges	Literature source
Qualitative	Grassroots forecasting (sales force composite)	Customer's intention to buy in the near future, understanding of and relationship with the customer	Laborious in retail situations and vague results	▪ Dalrymple, 1975 ▪ Klassen and Flores, 2001 ▪ Kahn, 2002
	Market research	Customer's preferences and insights, data for conjoint analyses, customer survey	Diverse data required to further improve results, need for identifying trends earlier and quantifying the value of marketing measures	▪ McFadden, 1986 ▪ Armstrong, 2001 ▪ Kahn, 2002 ▪ Wright, 2010 ▪ Shang *et al.*, 2015
	Expert forecast estimation	Domain knowledge, data foundation for decomposed decisions, quantitative analogies	Difficult to test the experts' hypotheses, no data-based decisions possible	▪ Sackman, 1974 ▪ Armstrong, 2001 ▪ Meyer and M., 2001 ▪ Wheelwright, and Hyndman, 2008
Quantitative	Time-series forecast	Historic sale figures, influence of different seasons, understanding of the development of a trend	Requires more recent data, needs to be interrupted when special events occur	▪ Hamilton, 1994 ▪ Pindyck and Rubinfeld, 1998 ▪ Smaros and Hellström, 2004 ▪ Makridakis *et al.*, 2008 ▪ Box *et al.*, 2011
	Causal demand forecast	Knowledge about the distinct factors influencing demand, large data sets, theoretical background, knowledge of causal relationships, segmentation strategies	Only a few causal relationships are known, and it is difficult to test them, requires recent data, expensive to produce forecasts	▪ Kahn, 2002 ▪ Armstrong *et al.*, 2006 ▪ Makridakis *et al.*, 2008

Table 1: Summary of forecasting methods

2.3 (Big) data analytics techniques

Big data is a term used to describe an exponentially growing mass of data. During the last 25 years, the amount of data in the world has sharply increased. In 2011, the International Data Corporation (IDC) estimated the total volume of data to equal 1.8ZB, equivalent to 1.8 billion terabytes (Chen *et al.*, 2014). It is forecasted to double at least every two years (Kambatla *et al.*, 2014). Accordingly, big data is generating tremendous attention worldwide these days (Wamba *et al.*, 2015; Akter and Wamba, 2016). The works of Wamba *et al.* (2017) and Gunasekaran *et al.* (2017) examined and illustrated the potential impact of big data on enterprises' performance. Generally, three characteristics—called the 3Vs and first defined by Doug Laney (2001)—are used to describe the term "big data": volume, variety, and velocity.

- *Volume* describes the vastness of data. It refers to not only large files measured in gigabytes and petabytes but also numerous transactions, files, and tables (Russom, 2011).
- *Variety* refers to the numerous different types of files and challenges of utilizing them. Whereas some data is structured, such as medical records, governmental statistics, and financial data, the more common and more challenging data formats are semi-structured, such as text, emails, and tweets, or unstructured, such as pictures and movies (Kambatla *et al*., 2014).
- *Velocity* has not gained as much attention as the other two Vs so far. It directly affects the value of data. Timely data capture and analysis ensure a dataset's value (Gantz and Reinsel, 2011). The more time that passes, the more irrelevant the data becomes (Zikopoulos *et al*., 2013). In supply chains, real-time monitoring of streaming data sources (e.g., radio-frequency identification (RFID) chips or cellular records) increases the possibilities for risk management or supply chain visibility (Sahay and Ranjan, 2008).

To analyze BDA's potential for demand forecasting, the findings of a narrative literature review are used to derive and cluster different analytics techniques. To the best of our knowledge, there is no single way to differentiate among analytics techniques. Therefore, we created a classification scheme distinguishing between computer models and self-service analytics. In order to classify techniques, we chose to separate them using an analytical approach. In total, five different techniques were classified: (1) data exploration, (2) advanced analytics, (3) interactive analysis and planning, (4) embedded analytics, and (5) stream analytics.

BDA techniques are embedded in the value chain of big data (Figure 2). The underlying scheme comprises the following stages: (i) *identify data sources*, (ii) *integrate data*, (iii) *analyze data*, and (iv) *produce actionable insights*. In general, insights should improve business management in the supply chain, and actionable insights create data in different volumes, varieties, and velocities. Similar process models have been constructed by consultancies that offer big data-related services and are comparable to those developed by Hagen *et al*. (2014) and Zikopoulos *et al*. (2013).[1]

[1] At this point, we should point out the difference between analytics and analysis. The two terms are often used interchangeably, yet their definitions deviate. Analysis is a process that extracts information and insights from

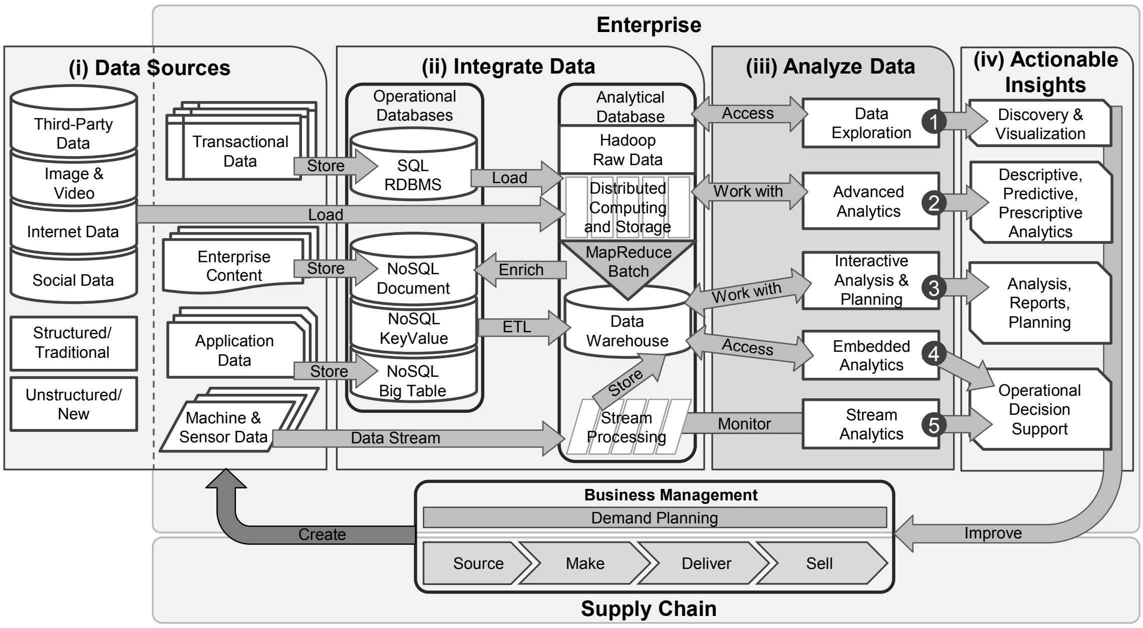

Figure 2: BDA techniques within the value chain of big data

2.3.1 Data exploration

Data exploration (EX) is a fundamental technique for understanding and exploiting a variety of big data. The goal of this technique is to produce insights that are relevant to businesses. It is a self-service analytics technique that is mostly performed directly by business users. The insights are discussed with other businesspeople to validate the outcomes and enrich discussions. Visualizations are a meaningful way to constructively steer these discussions. Companies understand the potential of analytics and empower business users to analyze the available data, a trend often called the "democratization of analytics" (Ramel, 2015). This trend highlights the movement in the domain of structured data. Business users can use self-service business intelligence (BI) to analyze data formerly restricted to senior management. However, business users might be overwhelmed by the variety of big data.

2.3.2 Advanced analytics

Advanced analytics (AA) addresses complex business questions, combining methods from statistics, data mining, and machine learning (Dorschel, 2015). AA is differentiated from simple analytics by the types of questions that can be answered with the computer model; simple analytics is associated with explicitly formulated questions that can be answered by a few SQL queries, while AA is data-driven and aims to reveal insights and answer implicit questions. The output of AA is computer models that can automatically process data, process

data. In contrast, analytics encompasses the methodology and art of analyzing. It is a sum of methods from statistics, data mining, and machine learning (Biswas & Sen, 2016). Technologies and tools that support the analytical process are also associated with the term "analytics" (Dorschel, 2015).

unstructured data, and merge data sets. Data scientists who perform AA tasks use data from various sources. In addition, AA is a broad field that can produce a wide variety of computer models (Wu *et al*., 2016). In order to enable the following analysis, in which the potential of AA to support forecasting methods is surveyed, AA is divided into (i) descriptive (DAA), (ii) predictive (PDAA), and (iii) prescriptive analytics (PSAA).

The first category, DAA, includes computer models that use clustering, association rules, and classifications. It also involves the process of preparing data sets. To be feasible inputs for computer models, diverse data sources need to be structured. Deep neural networks (Sainath, 2013) and text analytics (Osborne *et al*., 2013) are two of the most important types of data processing. Commonly applied DAA techniques are cluster analysis (Shmneli and Koppius, 2009), principle component analysis (Abbott, 2014), association rules (Kadochnikov, 2013), classification and decision trees (Abbott, 2014), ensemble learning (Ihler, 2012), logistic regression and neural networks (Kadochnikov, 2013), and naïve Bayes classifiers (Abbott, 2014). These techniques support handling of vast amounts of data.

The second category, PDAA, includes estimations. Estimation models are more difficult to create than classification models; whereas classification models have to predict only a few values (nominal, discrete target), continuous estimations must predict every value contained in the target variable (metric target) (Abbott, 2014). In addition to classical regression analysis, other methods are often applied, such as time-series models, regression trees, and artificial neural networks. The most appropriate methods for forecast demand, linear regression, and time-series analysis are described below. Artificial neural networks do not suit this endeavor because the model entails black-box elements.

The third category, PSAA, is a type of applied learning as it uses data obtained from descriptive, predictive, and domain knowledge and applies it to real cases. In this way, the system determines the best situation-specific action. PSAA is an emerging field involving advanced optimization and simulation (Puget, 2012) as well as game theory and decision analysis methods (Rouse, 2012).

2.3.3 Interactive analysis and planning

Interactive analysis and planning (IA&P) is related to BI and has undergone a fundamental shift from being an IT-led consolidation technique to an interactive analytics function for a wider range of business users (Sallam *et al*., 2015). These business users want to use analytics without elaborate IT or data science skills, and thus new BI initiatives focus on making analytics more accessible and pervasive in organizations. Within IA&P, only structured data from relational databases fit directly into the enterprise data warehouse (EDW). For extract transform load (ETL) jobs, the operational structured data and third-party data is merged into the EDW. IA&P solutions work directly with the EDW, and in combination with DAA and SA, more extracted data can be loaded into the EDW or similar processed analytical databases to enrich the data content. Data scientists spend around three-quarters of their time preparing the data for their models (Kadochnikov, 2013), and this time investment can be better utilized by giving business users access to processed data.

2.3.4 Embedded analytics

Embedded analytics (EA) focuses on condensing analytical capabilities to the point of impact. LaValle *et al*. (2010) highlight the importance of integrating analytics into business processes, stating that when companies start to embed information and insights into business processes, analytics becomes "alive." EA enables operations to make data-based decisions through automated and analytical processes. It can be defined as integrating analytical capabilities (i.e., reporting, dashboards, data discovery, predictive and prescriptive analytics) into business software for customer relationship management (CRM), enterprise resource planning (ERP), financial, and supply chain management applications (Aberdeen Group, 2014). According to a recent study (Dresner, 2013), embedding analytics into a business process is seen as an important task in the business community, while back-office functions (i.e., supply chain, IT, and operations) are regarded to have the most potential.

2.3.5 Stream analytics

The best practice for data in motion is stream analytics (SA). Streams of incoming data are instantaneously analyzed. An effective way of handling streaming data is called event processing (CEP). CEP works with predefined rules to analyze data in motion. With CEP, computers are programmed to count; filter; transform; alert after thresholds; detect correlated data, missing events, trends, erroneous data, or patterns; merge with data in a database; track entities; learn a model; and predict the next values (Perera, 2015). Computations must be performed quickly, and they need to be horizontally scalable. Often, data undergoes multiple computations performed by different nodes in a row.

Table 2 summarizes and compares the five main types of analytics in terms of data generation, storage and technologies, examples of analytics software, the roles of potential users, and actionable insights.

	Data generation	Data ingestion and storage technologies	Analytics software (examples)	Main users	Actionable insights	Literature sources
1) Data Exploration (EX)	Unstructured data: geospatial, temporal, text (social media, emails, newspapers, surveys, etc.), social networks	Analytics for the NoSQL database, distributed computing and storage, visualization technologies, cognitive computers, self-service data preparation, cloud computing	Data exploration and visualization: ▪ SAS Contextual Analytics ▪ Teradata Loom ▪ Oracle Endeca Information Discovery ▪ IBM Data Explorer, Watson Analytics Special analytical applications (e.g., Google Analytics)	Business users, business analysts, data scientist	Big data visualizations, data understanding, analysis, and reports	▪ Goetz and Lang, 2013 ▪ Evelson, 2015
2) Advanced Analytics (AA)	Every kind of data accessed from Hadoop, NoSQL DBs, or EDW	Data preparation technologies to bring structure to unstructured data, distributed computing and storage, Hadoop, graph analytics, cognitive computers, cloud computing	Advanced analytics modeling software: ▪ IBM SPSS Modeler ▪ SAS Enterprise Miner, Text Miner, SAS/OR, SAS Visual Data Discovery ▪ KNIME RapidMiner ▪ Oracle Data Miner	Data scientists	Descriptive, predictive, and prescriptive models; processed data; embedded analytics and stream analytics solutions	▪ Lozano, 2013 ▪ Kadochnikov, 2013 ▪ Laumanns and Squillante, 2013 ▪ Abbott, 2014
3) Interactive Analysis and Planning (IA&P)	Dimensional, structured data stored in the EDW (financial, CRM, etc.)	EDW, relational database, OLAP, SQL	Traditional BI software: ▪ SAS Office Analytics ▪ IBM Cognos ▪ QlikView ▪ SAP Lumira	Business users, business analysts	Financial analysis, reports, simple forecasts, interactive dashboards, and business understanding	▪ Petitclerc, 2013 ▪ Sallam *et al.*, 2015
4) Embedded Analytics (EA)	Processed data from machines, social media, CRM, temporal, geospatial, financial, etc.	Special-purpose solutions, cloud computing, technologies to embed analytics in CRM, ERP, etc.	Custom-built applications: ▪ Pentaho Embedded Analytics ▪ IBM MobileFirst for iOS	Operational user	Operational decision support, easy-to-grab data presentations, accessible on various devices, interactive computer models	▪ Logi Analytics, 2014 ▪ Aberdeen Group, 2014 ▪ Apple, 2014 ▪ Dorschel, 2015 ▪ Loshin, 2015
5) Stream Analytics (SA)	Incoming data streams from machines, the Internet (weather, market, etc.), and supply chains	Stream processing technology, complex event processing, fast-access database	Complex event processing software: ▪ IBM InfoSphere Streams ▪ Apache Spark and Storm ▪ Informatica Stream Analytics	Operational user	Event-driven notifications, dashboards, processed streaming data, real-time insights	▪ Aslett *et al.*, 2013 ▪ Spicer, 2013 ▪ Dorschel, 2015

Table 2: Overview of big data techniques

3 Interplay – the theoretical framing

After describing the conceptual elements of demand forecasting and BDA, the principles of the theoretical frame are discussed in order to explain the interplay between and performance of the defined elements.

3.1 Approach

According to the two-sided systems approach, an expository instantiation is created. Within this context, basic statements and justificatory knowledge are obtained from the literature. The objective is to develop a theoretical grounding to solve the design problem and address how BDA techniques could enhance demand forecasting methods. Now, the most relevant steps to connect the components are addressed. Following the information systems design approach proposed by Gregor and Jones (2007), the entities of interest are interconnected through decision cycles, which illustrate both a simplified sequence of steps and interlinked requirements between these steps (so-called “interplay”).

Because of the inherent uncertainty of decision-making and the factors influencing demand forecasting, we embed all elements in a contingency framework (Sousa and Voss, 2008). As the effectiveness and performance of managerial control systems is contingent upon the organizational structure, the arrangement of information in relation to technology and the environment reveals important principles and limitations for further implementation (Chenhall, 2003).

While “data” is colloquially used as a synonym for “information,” there is a distinction, which primarily lies in their function (Dubey *et al*., 2016). Represented by a set of values in the form of qualitative and quantitative variables, data is collected and analyzed. Once the data is processed and bound up in a certain context (e.g., a retail supply chain), information can be created by adding meaning and structure for the decision maker. Only through this process does the refined data become useful for businesses as information and suitable for supporting decision-making (Liew, 2007). In addition, there is a distinction in terms of a set of contingencies within the framework between the level of data and the level of information.

For problem-solving, the decision-making process is divided into the decision problem (“What?”) and its underlying method (“How?”). As illustrated in Figure 3, the starting point of the conceptual framework lies within the initial cycle between data and decision, represented by demand forecast. The overall iterative loop is closed by additional items in a counterclockwise direction as the “value chain of big data” is passed again.

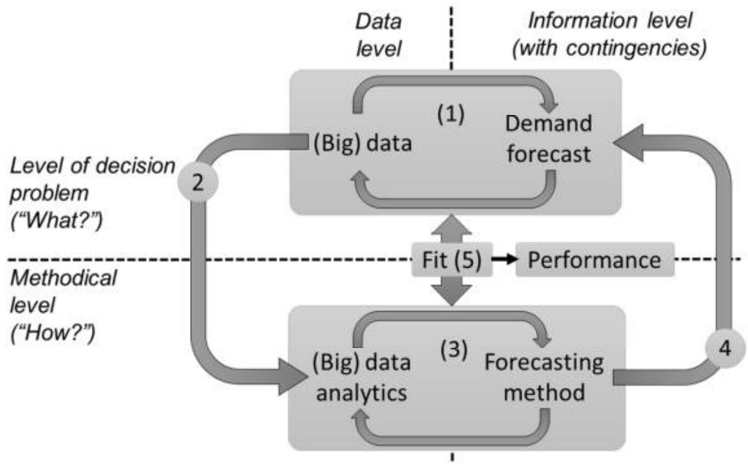

Figure 3: Approach to the link between BDA and demand forecasting

3.2 Interconnections

Decision-making problems (demand forecasts) are characterized by their scope and level of interaction (Madriakis *et al*., 2009). Different requirements and input variables are used if the focus is on long-term decisions (e.g., location of sites, mode of transport) or shorter-term decisions (e.g., category management, routing) (see 1 in Figure 3). Various types of data are created by almost countless supply chain activities. Thus, depending on the type of demand forecast, specific internal and external data sources have to be tapped (*(i) Data Sources*).

When business decisions are made, risk components have to be precisely measured and assessed (see 2 in Figure 3). Following this, forecasting methods serve as a ground for risk reduction within the decision-making process. Selection of the right method is therefore highly important. When historical data can be insufficient or not available, the method of choice is primarily dependent on the initial situation, and vice versa. Moreover, the tapped data has to be integrated (*(ii) Integrate Data*).

Every decision involves an element of uncertainty that cannot be measured or reduced with common statistical forecasting instruments. If the decision-maker is able to identify possible outcomes and subsequently estimate and assign probabilities, the state of uncertainty can be converted into a state of risk (Whalen and Bronn, 1988). BDA has these capabilities as its diverse applications take not only a number of variables into account but also a large volume, velocity, and variety of data. In order to improve demand forecasting, both the interaction and interoperability between the forecasting method and BDA applications have to be ensured (see 3 in Figure 3). Through this match, a computational shift is generated that not only improves but also expands the boundaries of demand forecasting. In this step, the real data analyses take place (*(iii) Analyze Data*).

BDA often requires additional data while adapting to the decision problem (see 4 in Figure 3) as the available information is usually incomplete and unreliable. Iteration within the loop stops until the model-driven decision support seems sufficient and the desired outcome is achieved (*(iv) Actionable Insights*).

Taking a closer look at the framework from the functional view of contingency theory proposed by Chenhall (2003), further limitations can be identified. Even when using a generic algorithm or a one-size-fits-all approach to demand forecasting, the relative impact of different drivers of demand, such as company size, industry, and product line, cannot be omitted. The items rely on adaption to evolving consumer behavior and market conditions. In order to increase performance (e.g., via improved business management), alignment in the form of internal consistency, congruency, or fit between these dimensions is required (see 5 in Figure 3). An effective organization has proper fit not only with the environment but also between its subsystems and technological components (Venkatraman, 1989, Van de Ven and Drazin, 1985). Thus, a central performance objective of BDA applications should be appropriate interaction with specific forecasting methods to achieve improved forecast accuracy.

3.3 Performance improvements

Accurate forecasting and demand planning is the basis of efficient supply chain management and execution (Barlas and Gunduz, 2001; Deschmukh and Mohan, 2016). The relations between the decision problem and methodical levels are therefore a decisive factor. Processing data with the wrong method results in false impressions, returns inaccurate forecasts, and changes everything else in the continuum, making the information gathered from the processed data unreliable and therefore impractical for decision-makers (Carbonneau *et al.*, 2008; Mello, 2015). While BDA, with its different techniques, adds factors to the methodical level, the role of fit becomes more apparent. In order to achieve improved forecast accuracy, BDA fits in between data and the forecasting method. It overcomes the performance gap by determining how to extract value from certain data to effectively support the decision-making process. After achieving performance fit by determining the "BDA method," coordination between data and forecasting method is unimportant. BDA extracts only the data needed to ensure that there is sufficient information for the forecasting decision. As a result, improved forecast accuracy is achieved (Hazen *et al.*, 2014).

Conventional methods also have limitations. Although historical data and the accumulated experience of decision-makers in the supply chain is traditionally used to support medium- and long-term decisions, it is almost impossible to improve supply chains and achieve a good level of accuracy (Heikkliä, 2002). The observed time frame of a forecast is the factor with the most influence in scenario planning techniques. The longer the predicted time horizon, the less accurate the forecast becomes (Schoemaker, 2006). In addition to this dependent variable, four aspects are specified to highlight the underlying challenge of traditional forecasting techniques in supply chains today. First, the available information is usually incomplete and unreliable as many parameters with direct effects on customer demand are omitted or not considered at all. In this regard, most samples are limited and do not match the parameters of demand planning concerning product dimension, geographical fit, or time horizon (Zhao *et al*., 2002). Second, complex algorithms generate simplistic statistical results because of statistical noise and unexplained variation in samples. Conventional approaches are often more appreciated as computations with shorter processing times and higher robustness can achieve the same level of accuracy as complex algorithms. Nevertheless, the underlying problem remains unsolved (Hazen *et al*., 2014). Third, the process of decision-making is not related to a body of knowledge. Often, past experiences are not captured in a system in the form of a learning curve (Towill, 1990). Patterns of speeding up and then slowing down within the dimensions of proficiency and experience are often not adequately considered (Meek *et al*., 2002). Fourth, qualitative instruments as well as quantitative methods require certain statistical skills and domain knowledge. As demand forecasting is still manually intensive, the results are highly dependent on the modelers' abilities and biases. Whether calculating order quantities or estimating additional safety stock, many demand planners prefer simplicity. If we assume that "the system does not know what the planner knows," it is much more likely that the planner cannot properly interpret the past (Hazen *et al*., 2014). As a complementary approach to traditional forecasting methods, to help further reduce uncertainty within the decision-making process, it would be worthwhile to fully automate the forecasting process in order to redesign the planners' work and provide knowledge instead of numbers.

4 Application – the exemplification by retail supply chains

In order to complete the two-sided approach, we now apply specific BDA techniques to practical retail cases in which they are used. As the benefits of BDA relate to improved demand forecast accuracy, the retail industry was chosen to identify, analyze, and assess BDA's practical value based on discussions with several retail enterprises located in Switzerland, which were our research partners. The retailer can be seen as the link between supply (consumer goods manufacturer) and demand (end-consumer) and is therefore challenged by constantly increasing volatility, uncertainty, complexity, and ambiguity (Randall *et al.*, 2011). To provide profound insight into BDA techniques' potential, we perform the following: (a) identification of typical forecast influence factors in retail supply chains, (b) assignment of BDA to these factors, (c) improvement of forecasting methods through BDA, and (d) final assessment based on different time horizons.

4.1 Typical forecast influence factors in retail supply chains

In general, non-retail professionals might view retail demand forecasting as straightforward time-series forecasts. However, these are only applicable in repetitive and short-term situations, and most retailers face varying product demand due to diverse factors, such as weather or short-lived trends. At a product's launch, retailers often possess no sales experiences but the product performs well due to promotions, the opening of new stores, or changing of the product assortment (Randall *et al.*, 2011; Sanchez-Rodrigues *et al.*, 2010). All of these determinants need to be identified and structured to acquire a profound understanding of demand challenges. According to the aggregated influencing factors in demand forecasting discussed by Souza (2014), we distinguished six demand forecast categories (Table 3).

Demand forecast categories	**Influence factors**	**Literature sources**
Products	Quality	▪ Hu *et al.*, 2008 ▪ Mankiw, 2011
	Competitors, substitutes, complements	▪ Mankiw, 2011
	Prices	▪ Herrmann and Huber, 2007 ▪ Mankiw, 2011
Consumer preferences	Fashions and trends	▪ Mankiw, 2011 ▪ Chatfield, 2013
	Buying behavior	▪ Hoyer and Brown, 1990
	Brand awareness and perception	▪ Hoyer and Brown, 1990
External factors	Weather	▪ Starr-McCluer, 2000 ▪ Murray *et al.*, 2010
	Special events	▪ Matheson, 2006
	Seasonality	▪ Mankiw, 2011 ▪ Chatfield, 2013
	Income level, economic outlook	▪ Mankiw, 2011
	Local development	▪ Baade and Dye, 1990
	Mega-trends (demographics, technology, climate etc.)	▪ Bowersox *et al.*, 2000 ▪ Mankiw, 2011
Marketing factors	Promotions	▪ Mankiw, 2011
	Advertisement	▪ Mankiw, 2011
Shop factors	Local competition	▪ Salmon and Tordjman, 2002
	Shop attractiveness	▪ Turley and Milliman, 2000
	Shop assortment and layout	▪ Turley and Milliman, 2000
Supply factors	Available products at point of sale	▪ Arvinder, 1998 ▪ Kurata, 2014
	Expiring products	▪ Huq *et al.*, 2005 ▪ Herbon, 2016

Table 3: Demand-influencing factors of retail supply chains

An essential task of demand forecasting is back-testing of the models to detect potential flaws and prevent their manifestation. In general, it can be established that diverse variables influence demand and, consequently, product sales. This suggest that the variables need to be considered during forecasting modeling in the context of the situation at hand.

4.2 Application of BDA to factors that influence forecasts

We now match BDA with the factors influencing demand for retail businesses. To achieve this aggregation, we followed several steps iteratively. Step 1 involves the establishment of a matrix in which the horizontals are BDAs and the verticals are factors influencing demand. Step 2 entails the determination of adequate time horizons for each field. During Step 3, each array of the matrix is be evaluated and assigned to support forecasting methods. Reasonable possible alternatives are suggested during Step 4. Finally, the padded fields of the matrices will be analyzed to reduce contradictions (i.e., fill empty fields) in Step 5. The authors tried to detect applications of BDA for every field but rejected examples that could not contribute to enhanced forecasts.

All the applications of BDA are listed in the following tables. Table 5 comprises applications for EX, IA&P, EA, and SA, and Table 6 contains all applications for AA. An indicator of the *time horizon* ranging from [s]hort- to [m]edium- to [l]ong-term exists for each application. However, an application can be used in more than one time frame. In these cases, the cell contains a time range, such as [S–M] for short- to medium-term. Next to the time horizon is an indicator, which shows the way in which generated insights can be used. The aim of analytics should be to improve current forecasting methods. The *numbers 1–5* indicate the forecasting methods explained above: [1] represents grassroots, [2] market research, [3] expert estimations, [4] time-series, and [5] causal forecast. Further, *quotation marks* are used as an indicator that the application relies on other aspects of analytics. For instance, DAA could evaluate social media trends that might influence demand, but to do so, prepared social media data originating from stream analytics is needed. Finally, some applications are not explicitly utilized to predict demand but might be useful (e.g., data scientists could create a price optimization tool with prescriptive analytics). Those applications are written in *italics.*

<table>
<tr><td>Time horizon</td><td colspan="2">S = Short-term</td><td colspan="2">M = Medium-term</td><td>L = Long-term</td></tr>
<tr><td>Supported forecasting method</td><td>1 = Grassroots</td><td>2 = Market research</td><td>3 = Expert estimations</td><td>4 = Time-series</td><td>5 = Causal forecasts</td></tr>
<tr><td>Other indications</td><td colspan="2">“Use cases in quotation marks rely</td><td colspan="3">Use cases written in italics do not</td></tr>
</table>

	on other analytics technique(s)”	*directly improve forecast*

Table 4: Explanation for BDA tables

		1) EX	3) IA&P	4) EA	5) SA
Products	Quality	Perform text analysis of product reviews and analysis of customer complaints [S–M,2]	*Analyze products' quality metrics*	*Shop managers can give immediate feedback about quality issues*	Not applicable
	Competitors, substitutes, complements	Text analysis of product reviews [M,2]	Not applicable	*“Recommend matching products or alternatives”*	Not applicable
	Prices	Visualize customers' sentiments about the price of new products [M,2,3]	Keep track of price information [M,1,2,5]	“Adjust prices with predicted influence” [S–M,1,5]	Track online prices of products [S–M,4,5]
Consumer preferences	Fashions and trends	Analyze the web traffic of an online shop to understand consumer demand [M,2]	Analyze how products' revenues develop over time [M,3]	“React to local buying trends by reordering more trending products” [S–M,1]	Not applicable
	Buying behavior	“Discover regional differences in shopping behavior” [M–L,2]	Analyze historic demand to predict future demand [S–M,1,3,4]	*“Prediction of visitor traffic to better schedule employees' working hours – prescription”* Information about predicted demand [S,1]	Evaluate shop purchases and detect unusual demand in real time [S,1,4,5]
	Brand awareness and perception	Analysis of social media sentiment about products and brands [M,2]	Analyze revenue for each brand and product [M,2,3,4]	*“Recommend products that match the customer's brand preferences”*	Track social media sentiments about products [S–M,2,4,5]
External factors	Weather	Not applicable	Not applicable	Be notified about weather conditions that influence demand [S,1]	Track weather conditions that influence demand [S,5]
	Special events	Keep track of important events [S–M,1,2]	Not applicable	Shop management indicates special local events and “order appropriate products” [S,1,4,5]	Not applicable
	Seasonality	Not applicable	Detect seasonality in the data [M,1,4]	Reaction to patterns [S–M, 4, 5]	Identification of regular intervals [S, 2, 4, 5]
	Income level, economic outlook	Text analysis of economic outlook predictions [M-L,3]	Not applicable	Not applicable	Not applicable
	Local development	Text analysis of newspapers about regional developments and visualize development using maps [M–L,3]	Not applicable	Not applicable	Not applicable

	Mega-trends (demographics, technology, climate etc.)	Text analysis of expert reports to understand the effect of climate on consumers via text analysis of scientific articles [L,3]	Not applicable	Detect patterns in the historic data that explain past effects of mega trends [L,3]	Not applicable
Marketing factors	Promotions	Not applicable	Analyze promotions in shops [S–M,1,2]	Local decisions about discounts "supported by predicted influence" [S,1,5]	Life update of promotions of each product in each shop [S,4,5]
	Advertisement	Analyze newspapers, magazines, billboards, etc. to detect running advertisements [M,2]	Not applicable	Not applicable	Keep track of running marketing campaigns on TV, the Internet, magazines, etc. [S–M,2,5]
Shop factors	Local competition	Screen data to detect new local competitors [M–L,3]	Analyze the financial and demand impact of competitors [M,2,3]	Not applicable	Not applicable
	Shop attractiveness	"Visualize a catchment area of shops on a map" [M,3]	*"Determine the ROI of renovations"*	Signal small local renovation work, outages, etc. [S,1,5]	Track social media sentiment about shops [S–M,1,2,5]
	Shop assortment and layout	"Visualize customer flow through the store" [M,3]	Analyze how the product assortment affects demand [M,4,5]	Indicate changes in local layout and assortment [S,1,5]	Not applicable
Availability factors	Available products at point of sale	Not applicable	Structured reports about stock excess and scarcity to "detect outliers" [S–M,1]	"Emergency reordering alert. Best next action when out of stock" [S,1]	Continuously updated stock of all goods, signals the need to restock [S,1,4,5]
	Expiring products	Not applicable	Analyze the value of wasted products [M,1]	*"Apply discounts in prescribed way"*	Track expiring products [S,1,4,5]

Table 5: Applying different analytics techniques to better forecast demand in retail supply chains

		2.1) DAA	2.2) PDAA	2.3) PSAA
Products	Quality	Assess product quality based upon reviews, customer feedback, warranty claims, etc. [M,1,2]	Predict influence from quality factors upon demand for a product [M,2,5]	*Optimize assortment with the best price-quality attributes*
	Competitors, substitutes, complements	Build a network of products to understand their interrelations [M,1,2]	Predict cross-price elasticity [M,2,5]	*Optimize the price structure of products*
	Prices	Structure the development of prices [M,2]	Predict price elasticity [M,2,5]	Not applicable
Consumer preferences	Fashions and trends	"Detect social media trends that could influence demand"; discover where trends start and how they spread; detect trend reversals [M,2]	Include social media trends in forecast models or test their influence on sales [M,2,5]	Simulate different product trend scenarios [M,3]
	Buying behavior	Segment customers according to their demand; detect purchase behavior with association rules; create a comprehensive view of individual customers from different data sources [M,2]	Predictive models based upon customer segments or individual customers; create forecast models with diverse input variables [S–M,2,5]; *predict customer visitor traffic*	Optimize demand forecast models with cost factors [S–M,1,4,5] *Optimize amount of cashiers based on visitor traffic; recommend products to customers*
	Brand awareness and perception	Describe brand awareness and perceptions [M,2]	Predict brand influence on product demand [M, 5]	Not applicable
External factors	Weather	*Categorize weather forecast data, but no direct support for the method*	Predict how the weather influences demand, "integrate categorized weather data into models" [S,2,4,5]	*Optimize assortment according to weather*
	Special events	Categorize events based on relevant criteria [S–M,1,2]	Predict how event type influences local demand [S–M,1,2,5]	*Optimize assortment depending on event type*
	Seasonality	Categorize products according to their seasonal demand [M,1,2]	Predict the influence of seasonality on demand [M,1,4,5]	*Optimize assortment depending on season*
	Income level, economic outlook	Categorize economic predictions, understand how experts correlate different indicators of economic development [M–L,3]	Predict the influence of the economy on demand [M,3,5]	*Optimize assortment depending on economic factors*
	Local development	Discover drivers of demand that arise from regional development, understand where customers come from [M–L,3]	Predict how much local and regional drivers influence demand [M–L,3,5]	*Find interesting new shop locations*
	Mega-trends (demographics, technology, climate. etc.)	Build a knowledge graph of related long-term trends and technologies to make future development more accessible [L,3]	Not applicable	Simulate trends' influence on consumer demand [M–L,3] *Optimize assortment*

Marketing factors	Promotions	Categorize promotion types [S–M,2]	Predict influence of promotions and discounts [S-M,2,5]	*Optimize the timing of promotions*, prescribe the amount of additional products needed [S,4,5]
	Advertisement	Categorize marketing campaigns, analyze social media reaction to campaigns [S–M,2]	"Predict demand influence of marketing campaigns" [S–M,2,5]	Prescribe how many more products shall be stocked [S–M,4,5]
Shop factors	Local competition	Create a network of local competitors, categorize new entrants and departures [M–L,3]	Adjust models according to changes in the categorization of competitors [M–L,3,5]	Simulate effect of new entrants on demand [M–L,3] *Find places to expand.*
	Shop attractiveness	Rate the attractiveness of the shops [M,2,3]	Integrate shop rating into model estimations [M,2,5]	*Calculate best time to invest in the shops' attractiveness*
	Shop assortment and layout	Categorize products depending upon their location in the shop, discover how much time customers spend in shops [M,1,2]	Change models according to shop assortments, determine influence of product location on demand [M,1,2,5]	*Optimize shop layout*
Supply factors	Available products at point of sale	Not applicable	Use prepared data about forecast errors to improve models, "adjust short time supply" [S–M,4,5]	Optimize a supply function based upon the cost of product scarcity and excess stock [S–M,1,4,5]
	Expiring products	Not applicable	Find drivers for wasted products [M,5]	Not applicable

Table 6: Applying advanced analytics to better forecast demand in retail supply chains

4.3 Improve forecasting methods with BDA

We analyze how BDA applications can improve forecasting methods. The bar chart (Figure 4) summarizes the applications presented in Tables 5 and 6. For the analysis, each application was counted once, with the total count indicating which types of BDA match the forecasting methods. This procedure has weaknesses since it does not consider how well (contextually) it supports the methods. We wanted to avoid consideration of additional factors that would have blurred our findings. Further research is required to analyze the fit in greater detail. Nevertheless, the results can be regarded as a first attempt to match analytics techniques to forecasting methods.

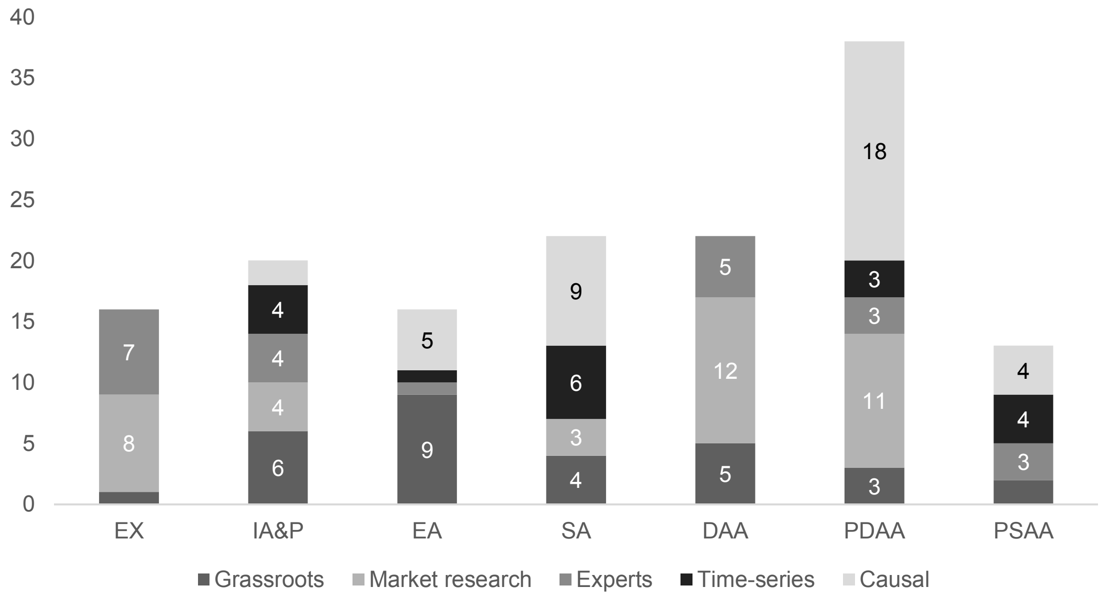

Figure 4: Improve forecasting methods in retail supply chains with BDA

Areas of improvement: Grassroots benefit most from EA, which could be labeled "smart grassroots forecasts." Market research benefits DAA, PDAA, and EX, enabling "data-rich market research." Structured expert forecasts benefit from EX, IA&P, and DAA and could therefore be named "smart and structured expert forecasts." All three qualitative forecasts benefit from quantitative analogies, which are more accessible with IA&P and thus become "interactive quantitative analogies." Time-series forecasts profit from SA and PSAA and become "rule based time-series." Moreover, causal forecasts benefit most from 1BDA with PDAA, SA, EA, and PSAA. This wide variety of data can be used for various forecast situations and thus becomes a "theory- and data-based causal forecast."

4.4 Improve forecast results with BDA

Given the possibility of improved forecasting methods, retailers might change their preferred method. Improvements in BDA could influence the forecasting methods in three ways. First, the improved forecasting method could *directly replace* the conventional methods. Second, the improvement of the method could *enable* it to be employed in a different situation. Third, the improvement could lead to a shift in methods in which an improved method *replaces* another method. Table 7 summarizes these changes. Most changes fall into the improvement category. Theory- and data-based causal forecasts can be used in nine situations. In the short run, it is a viable alternative to rule-based time-series. There is empirical evidence that the accuracy of forecasts can be improved when special events, such as forthcoming sales promotions, are considered to adjust the statistical forecast (Song *et al.*, 2011). Some retailers have implemented causal techniques, such as multiple linear regressions, to predict consumer demand. Chase (2014), for instance, used retail prices, media gross rating points, in-store merchandizing vehicles, sales promotions, and competitive retail activities to predict point-of-sale data. These findings are consistent with our suggestions. Retailers can test which method produces superior results or combine the results of different forecasts. To incorporate the influence of weather in forecasts, retailers should use theory- and data-based causal forecasts. Those replace time-series and market research because sufficient data can be gathered to predict the influence of weather with DAA and PDAA. In addition, causal forecasts can be applied to medium-term situations. The effect of price changes could be modeled when sufficient data on analog changes in the past becomes available. Moreover, complex time-series can be predicted more accurately with causal methods than with time-series. Blackburn *et al*. (2015) created a model to predict demand in the processing industry with PDAA that significantly outperformed time-series methods. Finally, BDA might improve long-term causal forecasts because the problem's segmentation is enhanced, more predictive data is available, and the forecast algorithm can be tested extensively.

Time horizon	Influence factor	Example situation or type of product	Likely forecast method without BDA	Change	Predicted forecasting method with BDA
Short	Constant demand	Toilet paper sales	Time-series	Improve Enable	Rule-based time-series T&DB causal forecast
	Seasonality	Fondue sales	Time-series	Improve Enable	Rule-based time-series T&DB causal forecast
	Simple continuous trend	Decreasing paper sales	Time-series	Improve Enable	Rule-based time-series T&DB causal forecast
	Promotions	20% discount on cookies	Time-series Market research	Replace Improve	T&DB causal forecast Data-rich market research
	Weather	Sunny weekend	Time-series Market research	Replace Replace	T&DB causal forecast
	Special events	Big concert nearby	Market research Grassroots	Improve Improve	Data-rich market research Smart grassroots
	Product quality issue	Lead discovered in noodles	Market research	Improve	Data-rich market research
Medium	Product price changes	Manufacturer demands higher prices	Market research	Improve Enable	Data-rich market research T&DB causal forecast
	New product iteration	New computer model released	Market research	Improve	Data-rich market research
	Change assortment	List unpopular products	Market research	Improve	Data-rich market research
	Customer preferences	New demand for gluten-free food	Market research	Improve	Data-rich market research
	Complex time-series	Car sales after new tax	Expert estimation Time-series	Improve Replace	S&S expert estimation T&DB causal forecast
	Shop renovation	New shop design	Expert estimation Grassroots	Improve Improve	S&S expert estimation Smart grassroots
Long	Economic outlook	Rising disposable income forecasted	Expert estimation Causal forecasts	Improve Improve	S&S expert estimation T&DB causal forecast
	New shop location	New shop in Geneva	Expert estimation	Improve	S&S expert estimation
	Local	New building	Expert estimation	Improve	S&S expert estimation

	development	complex nearby			
	Mega trends	Effect of climate change on car demand	Expert estimation	Improve	S&S expert estimation
S&S = smart and structured expert forecasts; T&DB = theory- and data-based causal forecast					

Table 7: Change in methods to forecast retail simulations

4.5 Usefulness of BDA over different time horizons

We will now evaluate the usefulness of BDA techniques regarding their application to forecasting over different time horizons. This additional analysis expands the understanding of BDA applications in the forecast setting. To elaborate which techniques are useful in which time horizon, all applications were counted for each time horizon. Table 7 summarizes the matching of techniques to time horizons. Regarding the results, it is striking that long-term forecasting has many fewer applications than short- and medium-term forecasting. *Long-term* forecasting can mainly benefit from EX and DAA, while the other techniques are less meaningful. This seems reasonable because SA focuses on real-time responses and data updates, whereas EA is designed for front-line workers, who are not faced with long-term, strategic decision-making. *Medium-term* forecasting can benefit mostly from PDAA and DAA and somewhat by EX and IA&P. On the one hand, AA generates profound insights in each demand-influencing area. On the other hand, SA's functionality is limited to applications for which real-time tracking generates a well-structured overview of the data. *Short-term* forecasting can mostly be improved by applying EA and SA. EA can empower shop managers in manifold ways and generate important data for short-term demand forecasts, and SA produces helpful data that could influence demand in the short term and keeps track of products in a store. The usefulness of BDA in retail supply chains for each time horizon is summarized in Figure 5.

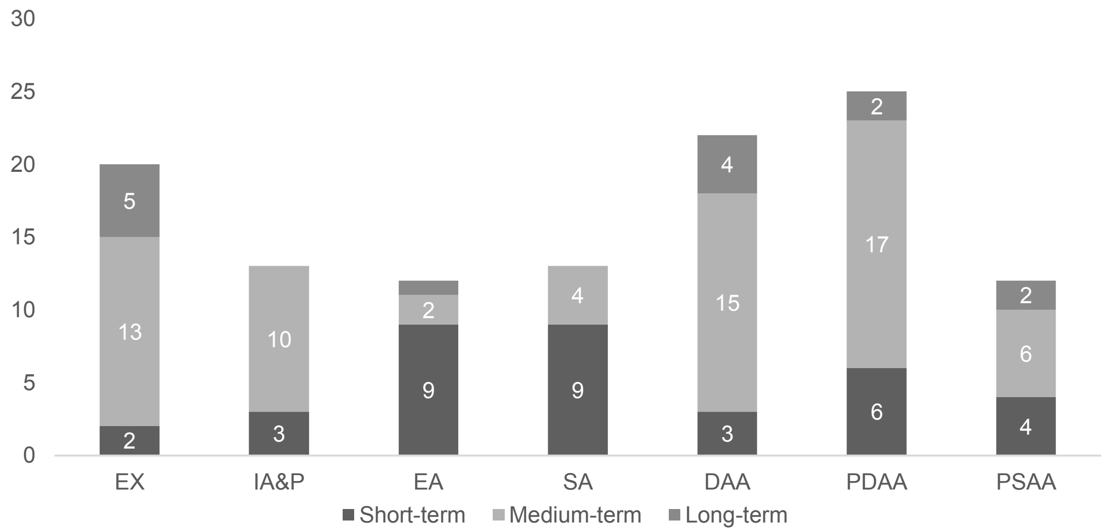

Figure 5: Usefulness of BDA in retail supply chains for each time horizon

5 Conclusion

5.1 Summary

This paper builds on the work of Blackburn *et al*. (2015) and Souza *et al*. (2014) by matching forecasting techniques with BDA. As a first proposition, it can be noted that big data in supply chains involves analytics of all kinds of data during the value creation process (Hazen *et al*., 2014). However, applying big data to analysis creates numerous challenges. To address the research question (How can BDA improve demand forecasting?), a two-sided approach was employed. According to this, a systematic overview is presented that provides both a theoretical frame for the interplay between elements and a conceptual framework demonstrating the existing match between data analytics and forecasting techniques. As a result, we obtained a more comprehensive understanding of the way in which to apply BDA within the decision-making process according to supply chains. Here, we demonstrated that the integration of different data sources is feasible. Combined with (retail) forecasts, the research analysis illustrated that, depending on the forecast situation, different analytics techniques can support the forecasting methods to overcome typical forecast challenges and improve the effectiveness of estimates. There is an existing relation between BDA and forecasting methods; the latter is clearly improved and was able to overcome previous challenges, such as forecast accuracy and consideration of many parameters with direct effects on customer demand, like different decision types and users (Blackburn *et al*., 2015).

The real case revealed that, in most situations, forecasting methods will benefit from BDA. An exemption are causal models created with AA, which introduced a new way of forecasting that will replace other methods. Causal models use recognized patterns and domain knowledge to make predictions as analyzing the past is always a necessary first step to discover hidden patterns. Besides causal models, the results suggest that situation-specific selection of forecasting methods will not fundamentally change. The applied method will change, however, due to additional data insights produced by BDA. The analysis also revealed that the different analytics techniques are used differently depending on the situation, indicating useful adoption of the contingency approach. Therefore, it is wise to define the desired forecast improvement first and then to plan potential BDA solutions.

Three main forecast situations that benefit from have been identified. First, strategic long-term estimation by corporate managers, leaders, or experts should use self-service EA and IA&P to support their decision-making with relevant insights. The solutions will benefit from prepared data that was initially processed by data scientists using DAA techniques. Second, tactical medium-term forecast situations should use DAA, PDAA, and EX to increase customer insights and trend awareness. Third, operational forecasts might benefit from improved time-series methods, including rules, which interrupt automatic forecasts when necessary. Operational forecasters can also employ PDAA to estimate demand with causal models. Such models were previously theory-led, but big data offers a vast data foundation to discover and quantify possible predictors. The appropriateness of forecasting methods for BDA depends significantly on the forecast situation. First, the target forecast improvement should be defined, and second, potential solutions with BDA should be specified. Nevertheless, data scientists are obliged to not over-fit models with complex algorithms.

Referring to the conceptual framework, it must be taken into account that certain components show industry-specific dependencies, meaning that the conducted elements can vary for different supply chains with specific objectives, time frames, and settings.

5.2 Machine learning as a next step?

According to our findings, BDA techniques are beneficial for companies and their supply chain activities (Wamba *et al*., 2015). The adoption of certain technologies is market- and time-specific. While conventional forecasting methods are still used within the market, innovative technologies can create further advantages. However, as competitive priorities change, it will become more important to analyze how customers behave than to analyze how much they buy. From this viewpoint, BDA has the potential to become a qualifying factor not only in the retail industry but also in any other supply chain. In particular, the growing adoption of the “Internet of Things” (IoT) lays a foundation within industries for further BDA applications (Deepa *et al*., 2016; Riggins and Wamba, 2015). Besides an innovative infrastructure, technologies such as the "blockchain," which uses a distributed ledger and decentralized database, will also change the way in which data is accessed and handled (Swan, 2015). According to the increasing role of data, integration of BDA within all kind of forecasting activities in the supply chain is inevitable (Khan, 2013).

In order to find an alternative approach, it seems to be necessary to take decision-making capabilities one step further (Wamba *et al.*, 2017; Tan *et al.*, 2015). Although various aspects of decision-making are being analyzed in statistics and related disciplines, such as dimension reduction, distributed optimization, or Monte Carlo sampling, machine learning can handle the growing complexity caused by proliferation of new data (Carbonneau *et al.*, 2008). Machine learning represents a shift from isolated data evaluation to a more sophisticated pattern recognition system by suitably modeling the time dependencies of available data. According to demand forecasting, a customer response is created through a model of demand signals that, in turn, is created by sensing market signals (Wong and Guo, 2010). The data-driven management planning approach does not only assist in situations where conventional methods fail but also identifies characteristics such as new market drivers. In particular, supervised and unsupervised learning techniques enable subroutines that automatically exclude redundant processes with below-average accuracy (Bumblauskas et al., 2016). Here, supervised learning is related to a multitude of data-mining techniques that transform given deterministic data based on the parameters of a forecast of target output data. On one hand, more demand variables are taken into account, but on the other hand, each value is weighted according to its significance. In addition to more accurate adaptation of different stock levels and reduction of planning efforts, decision-making processes in supply chain planning and design can benefit from supervised learning (Carbonneau *et al.*, 2008).

5.3 Limitations and outlook

Further research is required to analyze the use cases of BDA in relation to forecast situations to evaluate whether the matching is accurate. Such research would help to estimate the usefulness of diverse applications and support businesses in setting priorities. Moreover, an analysis of the costs of different BDA techniques would support businesses' investment decisions.

There are, of course, several limitations of this study. First, in addition to scientific articles, the literature review includes numerous non-academic sources from the information technology industry. However, they are essential to understand the techniques and technologies. To avoid bias, the sources were integrated with additional caution. Second, when analyzing the techniques' potential concerning forecast demand in retail supply chains, the authors limited the spectrum of demand forecast to primary demand, excluding secondary and tertiary demand, because the former often lets businesses directly determine the latter two. Third, the list of applications is not exhaustive; only applications that seemed meaningful or were encountered during the literature search were included. Fourth, the two analyses lack empirical proof because of the third limitation. Fifth, the usefulness of each application is assumed to be equal, but this does not reflect their actual usefulness.

References

Abbott, D. (2014). *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst.* Somerset, NJ, USA: Wiley and Sons.

Aberdeen Group (2014). *Embedded Analytics for the ISV: Supercharging Applications with BI.* Aberdeen Group.

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets*, *26*(2), 173-194.

Apple (2014). *Apple and IBM: The most enterprising apps ever*. Retrieved July 10, 2015, from Apple Homepage: https://www.apple.com/business/mobile-enterprise-apps/

Armstrong, J.S. (2001). *Principles of forecasting: A handbook for researcher and practitioners.* Springer Science & Business Media.

Armstrong, J.S., & Green, C.G. (2014). *Methodology Tree.* Retrieved from Forecasting Principles: http://www.forecastingprinciples.com/index.php/methodology-tree

Armstrong, J.S., Collopy, F., & Yokum, T. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting*, 22(3), 583-598.

Arvinder, P. (1998). Product distribution and service support strategy linkages: An empirical validation. *International Journal of Physical Distribution & Logistics Management, 28*(2), 143-162.

Aslett, M. (2013). *Big data reconsidered, separating hype from reality for hosting and cloud providers.* Retrieved from The 451 Group: https://451research.com/report-short?entityId=79479&referrer=marketing

Baade, R.A., & Dye, R. (1990). The impact of stadium and professional sports on metropolitan area development. *Growth and change*, *21*(2), 1-14.

Barlas, Y. & Gunduz, B. (2001). Demand forecasting and sharing strategies to reduce fluctuations and the bullwhip effect in supply chains. *Journal of the Operational Research Society March 2011, 62(3), 458-473.*

Bensinger, G. (2014). *Amazon Wants to Ship Your Package Before You Buy It.* Retrieved from Wall Street Journal Blog: http://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it/

Biswas, S., & Sen, J., (2016). A proposed Architecture for Big Driven Supply Chain Analytics. *International Journal of Supply Chain Management (IUP), June 2016.*

Blackburn, R., Lurz, K., Priese, B., Göb, R., & Darkow, I.-L. (2015). A predictive analytics approach for demand forecasting in the process industry. *International Transactions in Operational Research*, *22*(3), 407-428.

Bowersox, D.J., Closs, D.J., & Stank, T.P. (2000). Ten mega-trends that will revolutionize supply chain logistics. *Journal of Business Logistics,* 21(2), 1-15.

Box, G.E., Jenkins, G.M., & Reinsel, G.C. (2011). *Time-series analysis: Forecasting and control.* John Wiley & Sons.

Bright, J.R. (1978). Practical Technology Forecasting. Austin: The Industrial Management Center, Inc

Bumblauskas, D., Nold, H., Bumblauskas, P., Igou, A. (2017). Big data analytics: trasforming data to action. Business Process Management Journal, 23(3), 703-720.

Carbonneau, R.; Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184(3), 1140-1154.*

Chambers, J.C., Mullick, S.K., & Smith, D.D. (1971). How to Choose the Right Forecasting Technique. *Harvard Business Review*, 45(1), 45-71.

Chase, C. (2014). *Intermittent Demand Forecasting and Multi-tiered Causal Analysis.* North Carolina: SAS Institute Inc.

Chatfield, C. (2013). *The analysis of time series: An introduction.* CRC press.

Che, H., Chen, X., & Chen, Y. (2012). Investigating Effects of Out-of-Stock on Consumer Stockkeeping Unit Choice. *Journal of Marketing Research,* 49(4), 502-513.

Chen, D., Preston, D., & Morgan, S. (2015). How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management. *Journal of Management Information Systems,* 32(4), 4-39.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications,* 19(2), 171-209.

Chenhall, R.H. (2003). Management control systems design within its organizational context: findings from contingency-based research and directions for the future. *Accounting, Organizations & Society,* 28(2-3), 127-168.

Christopher M. & Peck, H. (2003): Building the resilient supply chain. *The International Journal of Logistics Management,* 15(2), 1-14.

Clarke, L. (2013). *Analysis: How Tesco and Otto are using data to forecast demand.* Retrieved from RetailWeek: http://www.retail-week.com/topics/analysis-how-tesco-and-otto-are-using-data-to-forecast-demand/5053784.article

Croxton, K. L., Lambert, D. M., García-Dastugue, S. J., & Rogers, D. S. (2002). The demand management process. *The International Journal of Logistics Management*, 13(2), 51-66.

Dalal, S. & Athavale, V. (2012). Analysing Supply Chain Strategy Using Case-Based Reasoning. *Journal of Supply Chain Management Systems,* 1(3), 40-48.

Dalrymple, D.J. (1975). Sales forecasting methods and accuracy. *Business Horizons*, 18(6), 69-73.

Dekimpe, M.G., & Hanssens, D.M. (1995). The persistence of marketing effects on sales. *Marketing Science*, *14*(1), 1-21.

Deschmukh, A. & Mohan, A. (2016). Demand Chain Management: The Marketing and Supply Chain Interface Redefined. *Journal of Supply Chain Management,* 13(1), 20-36.

Dorschel, J. (2015). *Praxishandbuch Big Data.* Karlsruhe, Germany: Springer Gabler.

Dresner, H. (2013). *Wisdom of Crowds Embedded Business Intelligence Market Study.* Dresner Advisory Services, LLC.

Dubey, R., Gunasekaran, A., Childe, S. J., Wamba, S. F., & Papadopoulos, T. (2016). The impact of big data on world-class sustainable manufacturing. *The International Journal of Advanced Manufacturing Technology*, 84(1-4), 631-645.

Duhigg, C. (2012). Companies Learn Your Secrets. *The New York Times Magazine*, 16, 2012.

Evelson, B. (2015). *The Forrester Wave: Enterprise Business Intelligence Platforms, Q1 2015,* Cambridge: Forrester.

Fernandes, R., Pinho, C., & Gouveia, B. (2015). Supply chain networks design and transfer-pricing. *The International Journal of Logistics Management*, 26(1), 128-146.

Fernie, J., & Grant, D. B. (2008). On-shelf availability: The case of a UK grocery retailer. *The International Journal of Logistics Management*, 19(3), 293-308.

Flynn, B., Koufteros, X. & Guanyi, L. (2016). On Theory in Supply Chain Uncertainty and its Implications for Supply Chain Integration. *Journal of Supply Chain Management,* 52(3), 3-27.

Galasso, F., Merce, C. & Grabot, B. (2009). Decision support framework for supply chain planning with flexible demand. *International Journal of Production Research,* 47(2), 455-478.

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC iview*, p. 1142.

Gnanendran, K., & Iacocca, K. (2015). The point of purchase decision in a supply chain with value-added reselling. *International Journal of Production Research, 53*(22), 6689-6700.

Goetz, T.W., & Lang, A. (2013). *Case Study: Social Media Analytics.* New York: IBM.

Gregor, S. & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems*, 8(5), 312-335.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S.F., Childe, S.J., Hazen, B., Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research,* 70, 308-317.

Hämäläinen A. (2011): Effect of uncertainties on demand forecasting. *Conference Paper the 2011 World Congress in Computer Science Computer Engineering and Applied Computing.*

Hamilton, J.D. (1994). *Time series analysis.* Princton: Princeton University Press.

Hanson, B., Hodgson, T., Kay, M., King, R., & Thoney-Barletta, K. (2015). On the economic lot scheduling problem: stock-out prevention and system feasibility. *International Journal of Production Research, 53*(16), 4903-4916.

Hazen, B.T., Boone, C.A., Ezell, J.D., & Jones-Farmer, L.A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72-80.

Hazen, B.T., Boone, C.A., Skipper, J.B., Hill, R.R. (2016). Back in business: Operations research in support of big data analytics for operations and supply chain management *Annals of Operations Research*, p. 1-22.

Hazen, B.T., Skipper, J.B., Ezell, J.D., & Boone, C.A. (2016). Big Data and predictive analytics for supply chain sustainability: A theory-driven research agenda. *Computers & Industrial Engineering*.

Heikkilä, J. (2002). From supply to demand chain management: efficiency and customer satisfaction. *Journal of Operations Management,* 20(6), 747-767.

Herbon, A. (2016). Optimal piecewise-constant price under heterogeneous sensitivity to product freshness. *International Journal of Production Research, 54*(2), 365-385.

Herrmann, A.X., & Huber, F. (2007). The influence of price fairness on customer satisfaction: an empirical test in the context of automobile purchases. *Journal of Product & Brand Management*, 16(1), 49-58.

Hevner, A.R., March, S.T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 23*(1), 75-105.

Horvath, L. (2001). Collaboration: The key to value creation in supply chain management. *Supply Chain Management: An International Journal,* 6(5), 205-207.

Hoyer, W. D., & Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *Journal of Consumer Research*, *17*(1), 141-148.

Hu, N., Liu, L., & Zhang, J.J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, *9*(3), 201-214.

Huq, F., Asnani, S., Jones, V., & Cutright, K. (2005). Modeling the influence of multiple expiration dates on revenue generation in the supply chain. *International Journal of Physical Distribution & Logistics Management, 35*(3), 152-160.

Kadochnikov, N. (2013). Data Mining Basics. Business Analytics Transformation IBM.

Kahn, K.B. (2002). An exploratory investigation of new product forecasting practices. *Journal of Product Innovation Management*, *19*(1), 133-143.

Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), 2561-2573.

Khan, R. (2013). Business Analytics and Supply Chain Performance: An Empirical Perspective, *International Journal of Operations and Logistics Management,* 2(3), 43-56.

Klassen, R.D., & Flores, B.E. (2001). Forecasting practices of Canadian firms: Survey results and comparisons. *International Journal of Production Economics*, *70*(2), 163-174.

Kurata, H. (2014). How does inventory pooling work when product availability influences customers' purchasing decisions? *International Journal of Production Research, 52*(22), 6739-6759.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note 6*, Gartner, Stamford, CT, USA, White Paper, 2001.

Laumanns, M., & Squillante, M. (2013). *Analytics Education: Stochastic and Robust Optimization (Optimization under Uncertainty).* New York: IBM Corporation.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2010, December 21). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 20-32.

Liew, A. (2007). Understanding Data, Information, Knowledge and their Inter-Relationships. *Journal of Knowledge Management Practice,* 8(2), 1-16.

Logi Analytics. (2014). *The complete guide to Embedded Analytics.* Retrieved from Logi Analytics: http://images.learn.logixml.com/Web/LogiAnalyticsInc/%7B9479a99f-f266-4553-a981-2c2181ed65ea%7D_LogiAnalytics-eBook-02112014.pdf

Loshin, D. (2015). *Integrated analytics gives users operational intelligence edge*. Retrieved from Techtarget: http://searchbusinessanalytics.techtarget.com/feature/Integrated-analytics-gives-users-operational-intelligence-edge

Lowson, R. (2001). Retail Operational Strategies in Complex Supply Chains. *International Journal of Logistics Management,* 12(1), 97-111.

Lozano, A.C. (2013). *Analytics Education: Advanced Data Mining.* New York: IBM Corporation.

Lundholm, R.J., McVay, S.E., & Randall., T. (2010). Forecasting sales: A model and some evidence from the retail industry. *Working paper, http://ssrn.com/abstract=486162.*

Makridakis, S., Hogarth, R. & Gaba, A. (2009). Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting,* 25(4), 794-812.

Makridakis, S., Wheelwright, S.C., & Hyndman, R.J. (2008). *Forecasting Methods and Applications, Third Edition.* John Wiley & Sons.

Mankiw, G.N. (2011). *Brief Principles of Macroeconomics.* Cengage Learning.

Martino, J.P. (1972). Forecasting The Progress Of Technology. In J.P.Martino (Ed.), An Introduction to Technological Forecasting (2nd ed., pp. 13-23). New York: Gordon and Breach Science Publishers.

Matheson, V. (2006). Mega-Events: The effect of the world's biggest sporting events on local, regional, and national economies. *Economics Department Working Papers. Paper 68*.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science*, *5*(4), 275-297.

Meek, C., Thiesson, B., & Heckerman, D. (2002). The Learning-Curve Sampling Method Applied to Model-Based Clustering. *Journal of Machine Learning Research, 2(3), 397.*

Mell, J. (2015). Internal and External Collaboration: The Keys to Demand-Supply Integration. *Foresight: The International Journal of Applied Forecasting,* (36), 16-20.

Meredith, J. (1993). Theory Building through Conceptual Methods. *International Journal of Operations & Production Management,* 13(5), 3-11.

Mest, D.P. & Plummer (2003). Analysts' rationality and forecast bias: Evidence from sales forecasts. *Review of Quantitative Finance Accounting,* 21(2), 103-122.

Meyer, M.A., & Booker, J.M. (2001). *Eliciting and analyzing expert judgment a pracitcal guide.* SIAM.

Mishra, D., Gunasekaran, A., Childe, S., Papadopoulos, T., Dubey, R., & Wamba, S. (2016). Vision, applications and future challenges of Internet of Things: a bibliometric study of the recent literature. *Industrial Management & Data Systems,* 116(7), 1331-1355.

Murray, K.B., Di Muro, F., Finn, A., & Leszczyc, P. (2010). The effect of weather on consumer spending. *Journal of Retailing and Consumer Services*, 17(1), 512-520.

Nikolopoulus, K., Babai, Z., & Bozos, K. (2016). Forecasting supply chain sporadic demand with nearest neighbor approaches. *International Journal of Production Economics,* 177, 139-148.

Osborne, M., Plumley, J., & Zaratsian, D. (2013). *Seeing the Voice of the Customer: Identifying Root Cause with Text Analysis Visualization.* SAS Institute Inc. World Headquarters.

Perera, S. (2015). *Scalable Realtime Analytics with declarative SQL like Complex Event Processing Scripts.* Retrieved from Slideshare.net: http://www.slideshare.net/hemapani/scalable-realtime-analytics-with-declarative-sql-like-complex-event-processing-scripts

Petitclerc, M. (2013). *Introduction to OLAP & Multidimensional Analysis.* New York: IBM.

Pindyck, R.S., & Rubinfeld, D.L. (1998). *Econometric models and economic forecasts.* Boston: Irwin/McGraw-Hill.

Ramel, D. (2015). Skills Shortage Prompts Democratization of Big Data Analytics. *Enterprise Systems Journal*. Retrieved from https://esj.com/articles/2015/02/09/big-data-democratization.aspx

Randall, S., Gibson, B.J., Defee, C.C., Williams, B.D. (2011). Retail supply chain management: key priorities and practices. *The International Journal of Logistics Management,* 22(3), 390-402.

Richard L.P., & Morgan S. (2012). A Demand-side Perspective on Supply Chain Management. *Journal of Supply Chain Management,* 48(2), 7-13.

Riggins, F.J., & Wamba, S.F. (2015). Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics. In : *System Sciences (HICSS), 2015 48th Hawaii International Conference*, pp. 1531-1540.

Russom, P.T. (2011). *Big data analytics.* TDWI Best Practices Report.

Sackman, H. (1974). *Delphi assessment: Expert opinion, forecasting, and group proecess.* Santa Monica: Rand Corporation.

Sahay, B.S. & Ranjan, J. (2008) "Real Time Business Intelligence in Supply Chain Analytics". *Information Management & Computer Security,* 16 (1), 28-48.

Sainath, T.N. (2013). *Deep Neural Networks.* Speech and Language Algorithms Group, IBM.

Sallam, R.L., Hostmann, B., Schlegel, K., Tapadinhas, J., Parenteau, J., & Oestreich, T.W. (2015). *Magic Quadrant for Business Intelligence and Analytics Platforms.* Gartner.

Salmon, W.J., & Tordjman, A. (2002). The Internationalisation of Retailing. *Retailing: Comparative and International Retailing*, 4, 84.

Sanchez-Rodrigues, V., Potter, A., Naim, M.M. (2010). Evaluating the causes of uncertainty in logistics operations. *The International Journal of Logistics Management*, 21(1), 45-64.

Schoemaker, P.J.H. (1993). Multiple Scenario Development: Its Conceptual and Behavioral Foundation, *Strategic Management Journal,* 14(3), 193-213.

Shang, G., Saladin, B., Fry, T., & Donohue, J. (2015). Twenty-six years of operations management research (1985–2010): authorship patterns and research constituents in eleven top rated journals. *International Journal of Production Research,* 53(20), 6161-6197.

Shmueli, G., & Koppius, O. (2009). The Challenge of Prediction in Information Systems Research. *Robert H. Smith School Research Paper*, 6-152.

Silva, A., & Reilly, S. (2014). A Big Data Methodology for Bridging Quantitative and Qualitative Political Science Research. *Conference Papers – American Political Science Association,* 1/1/2014, 1-24.

Simon, H.A. (1996). *The Sciences of the Artificial.* Cambridge: MIT Press.

Sinha, A., Millhiser, W.P., He, Y. (2016). Matching supply with demand in supply chain management eductaion. *The International Journal of Logistics Management*, 27(3), 837-861.

Smaros, J., & Hellström, M. (2004). Using the assortment forecasting method to enable sales force involvement in forecasting: A case study. *International Journal of Physical Distribution & Logistics Management,* 34(2), 140-157.

Song, H., Gao, B.Z., & Lin, V.S. (2011). Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system. *Journal of Forecasting,* 29(2), 295-310.

Sousa, R. & Voss, C.A. (2008). Contingency research in operations management practices. Journal of Operations Management, 26(6), 697-713.

Souza, G.C. (2014). Supply chain analytics, *Business Horizons*, 5(57), 595-605.

Spicer, M. (2013). *Capture and Analyze Data in Motion with IBM InfoSphere.* New York: IBM Corporation.

Starr-McCluer, M. (2000). *The Effects of Weather on Retail Sales.* Federal Reserve Board of Governors.

Swan, M. (2015). Blockchain : blueprint for a new economy. Beijing O'Reilly 2015

Tan, K.T., Than, Y., Ji, G., Ye, F., Chang, D. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastruture based in deduction graph. *International Journal Production Economics*, 165, 223-233.

Towill, D. (1990). Forecasting learning curves. *International Journal of Forecasting,* 6(1), 25-38.

Turley, L. W., & Milliman, R. E. (2000). Atmospheric effects on shopping behavior: a review of the experimental evidence. *Journal of Business Research*, 49(1), 193-211.

Van den Ven, A.H., & Drazin, R. (1985). Alternative Forms of Fit in Contingency Theory. *Administrative Science Quarterly,* 30(4), 514-539.

Venkatraman, N. (1989). The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence. *The Academy of Management Review,* 14(3), 423-444.

Wamba, S.F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246.

Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Rameshwar, D., Childe, S.J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Buisness Research*, 70, 356-365.

Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

Whalen, T. & Bronn, C. (1988). Essentials of Decision Making Under Generalized Uncertainty. Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making. *Volume 310 of the series Lecture Notes in Economics and Mathematical Systems,* pp. 26-47.

Whetten, D.A. (1989). What constitutes a theoretical contribution?. *Academy of Management Review*, 14(4), 490-495.

Wong, C., Lai, K., & Cheng, T.C.E. (2011). Value of Information Integration to Supply Chain Management: Roles of Internal and External Contingencies. *Journal of Management Information Systems,* 28(3), 161-200.

Wong, W. K., & Guo, Z. X. (2010). A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm. *International Journal of Production Economics*, *128*(2), 614-624.

Wright, T.C. (2010). *Real Life Examples of Qualitative Forecasting*. Retrieved from Small Business Chron: http://smallbusiness.chron.com/real-life-examples-qualitative-forecasting-72990.html

Wu, L., Yue, X., Jin, A., Yen, D.C. (2016). Smart supply chain management: a review and implications for future research. *The International Journal of Logistics Managemenet*, 27(2), 395-417

Yazdanparast, A., Manuj, I., Swartz, S.M. (2010). Co-creating logistics value: a service-dominant logic perspective. *The International Journal of Logistics Managemenet*, 21(3), 375-403

Zhan, Y., Tan. K.H., Li, Y., Tse, Y.K. (2016). Unlocking the power of the big data in new product development. *Annals of Operations Research,* pp. 1-19.

Zhao, X., Xie, J., & Leung, J. (2002). The impact of forecasting model selection on the value of information sharing in a supply chain. *European Journal of Operational Research,* 142(2), 321-344.

Zikopoulos, P.C., de Roos, D., Parasuraman, K., Deutsch, T., Corrigan, D., & Giles, J. (2013). *Harness the power of big data.* The McGraw-Hill Companies.