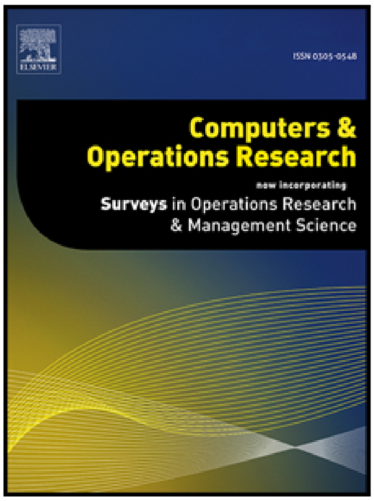
Accepted Manuscript



Big data analytics in supply chain management: A state-of-the-art literature review

Truong Nguyen , Li ZHOU , Virginia Spiegler ,
Petros Ieromonachou , Yong Lin

PII: S0305-0548(17)30168-5
DOI: 10.1016/j.cor.2017.07.004
Reference: CAOR 4285

To appear in: *Computers and Operations Research*

Received date: 1 December 2016
Revised date: 30 June 2017
Accepted date: 3 July 2017

Please cite this article as: Truong Nguyen , Li ZHOU , Virginia Spiegler , Petros Ieromonachou , Yong Lin , Big data analytics in supply chain management: A state-of-the-art literature review, *Computers and Operations Research* (2017), doi: 10.1016/j.cor.2017.07.004

Big data analytics in supply chain management:

A state-of-the-art literature review

Truong Nguyen(a); Li ZHOU (a); Virginia Spiegler (b); Petros Ieromonachou (a); Yong Lin (a)

a: Connected Cities Research Group, Systems Management and Strategy Department, Faculty of Business, University of Greenwich, London, UK

Corresponding author Li Zhou email:

ZL14@gre.ac.uk<mailto:ZL14@gre.ac.uk>

b: Kent business school, University of Kent, Kent, UK

Abstract:

The rapidly growing interest from both academics and practitioners in the application of big data analytics (BDA) in supply chain management (SCM) has urged the need for review of up-to-date research development in order to develop a new agenda. This review responds to the call by proposing a novel classification framework that provides a full picture of current literature on where and how BDA has been applied within the SCM context. The classification framework is structurally based on the content analysis method of Mayring (2008), addressing four research questions: (1) in what areas of SCM is BDA being applied? (2) At what level of analytics is BDA used in these SCM areas? (3) What types of BDA models are used in SCM? (4) What BDA techniques are employed to develop these models? The discussion tackling these four questions reveals a number of research gaps, which leads to future research directions.

Keywords: Literature review; Big data; Big data analytics; Supply chain management; Research directions.

1. Introduction

With the fast-paced and far-reaching development of information and communication technologies (ICTs), big data (BD) has become an asset for organizations. BD has been characterized by 5Vs: volume, variety, velocity, veracity, and value (Wamba et al., 2015; Assunção et al., 2015; Emani et al., 2015). Volume refers to the magnitude of data, which has exponentially increased, posing a challenge to the capacity of existing storage devices (Chen and Zhang, 2014). Variety refers to the fact that data can be generated from heterogeneous sources, for example sensors, Internet of things (IoT), mobile devices, online social networks, etc., in structured, semi-structured, and unstructured formats (Tan et al., 2015). Velocity refers to the speed of data generation and delivery, which can be processed in batch, real-time, nearly real-time, or streamlines (Assunção et al., 2015). Veracity stresses the importance of data quality and level of trust due to the concern that many data sources (e.g. social networking sites) inherently contain a certain degree of uncertainty and unreliability (Gandomi and Haider, 2015; IBM, 2012; White, 2012). Finally, Value refers to the process of revealing underexploited values from BD to support decision-making (IDC, 2012; Oracle, 2012).

Among those 5Vs, veracity and value, which represent the rigorousness of Big Data Analytics (BDA), are particularly important because without data analysis, other BD processing aspects such as collection, storage, and management would not create much value (Huang et al., 2015; Chen and Zhang, 2014; Babiceanu and Seker, 2016).

BDA involves the use of advanced analytics techniques to extract valuable knowledge from vast amounts of data, facilitating data-driven decision-making (Tsai et al., 2015). Supply chain management (SCM) has been extensively applying a large variety of technologies, such as sensors, barcodes, RFID, IoT, etc. to integrate and coordinate every linkage of the chain. Therefore, not surprisingly, supply chains (SCs) have been revolutionized by BDA and its application in SCM has been reported in a number of special issues (Wamba et al., 2015; Gunasekaran et al., 2016; Wamba et al., 2017). Indeed, BDA is reported to be an emerging SC game changer (Fawcett and Waller, 2014; Dubey et al., 2016), enabling companies to excel in the current fast-paced and ever-changing market environment. Empirical evidence demonstrates multiple advantages of BDA in SCM including reduced operational costs, improved SC agility, and increased customer satisfaction (Sheffi, 2015; Ramanathan et al., 2017) and consequently, there is an increasing interest in identifying a specific skill set for SCM data scientists (Waller and Fawcett, 2013; Schoenherr and Speier-Pero, 2015). Although the expectation of BDA adoption to enhance SC performance is rather high, a recent report found that only 17% of enterprises have implemented BDA in one or more SC function (Wang et al., 2016). The main reasons for low uptake are the lack of understanding of how it can be implemented, the inability to identify suitable data (Schoenherr and Speier-Pero, 2015), low acceptance, routinization and assimilation of BDA by organizations and SC partners (Gunasekaran et al., 2017), and data security issues (Fawcett and Waller, 2014; Dubey et al., 2016). This *motivates* our exploration of the existing research and the applications of BDA in SCM.

There are a number of literature reviews of BDA applications in the SCM context, but most of them tend to focus on a specific operational function of the SC. For instance, O’Donovan et al. (2015), Dutta and Bose (2015), and Babiceanu and Seker (2016) conducted literature reviews on material flow in manufacturing operations while Wamba et al. (2015) focused on logistics applications. A literature review that takes a broad perspective of SC as a whole and cross-maps with BDA techniques in SCM is yet scarce (Olson, 2015; Addo-Tenkorang and Helo, 2016; Hazen et al., 2016; Wang et al., 2016; Mishra et al., 2016). Our literature review develops a classification framework, which identifies and connects SC functions with levels of analytics, BDA models and techniques. Our review scope aims to provide a full picture of where and how BDA has been applied in SCM, by mapping BDA models and techniques to SC functions.

To obtain the objective, the literature review attempts to address the following four research questions:

(1) In what areas of SCM is BDA being applied?

(2) At what level of analytics is BDA used in these SCM areas?

(3) What types of BDA models are used in SCM?

(4) What BDA techniques are employed to develop these models?

The paper is structured as follows. Section 2 describes the review methodology used for the literature search and delimitation. It develops the classification framework for this review. Section 3 undertakes a review in line with the developed framework. Section 4 discusses the findings. Section 5 provides the avenues for future research. Section 6 concludes the review and presents the research limitations.

2. Review methodology

To address the research questions, the review methodology is based on the content analysis approach proposed by Mayring (2008). This approach has been adopted by a number of highly cited review papers in SCM literature, such as Seuring and Muller (2008), Seuring (2013), and Govindan et al. (2015). In particular, the review is systemically conducted in accordance with the four-step iterative process:

- Step 1: Material collection, which entails a structured process of search and delimitation of articles.
- Step 2: Descriptive analysis, which provides general characteristics of the studied literature.
- Step 3: Category selection, which aims to construct a classification framework based on a set of structural dimensions and analytic categories.
- Step 4: Material evaluation, which analyses articles based on the proposed classification framework and interprets the results.

2.1 Material collection

Before searching for articles, it is essential to identify an effective set of keywords that can capture the synthesis of existing literature related to our research topic. We classified keywords into two groups:

- Group 1: Words related to BDA: "Big Data"; "Data analytics"; "Data mining"; "Machine learning"; "Descriptive analytics", "Predictive analytics", "Prescriptive analytics".
- Group 2: Words related to SCM: "Supply Chain"; "Purchasing"; "Procurement"; "Manufacturing"; "Inventory"; "Storage assignment"; "Order picking"; "Logistics"; "Transport".

Note that "inventory", "storage assignment", and "order picking" indicate three major functions of warehousing operations in SCM. The reason we used these specific terms rather than "warehouse" is to avoid the confusion with "data warehousing", a well-established, technical research topic of BDA.

The search was conducted based on all possible pairs between those two types of keywords within the timeline from 2011 to the first half of 2017 on well-known academic databases, i.e. Science Direct, Emeralds, Scopus, EBSCO, and IEEE Xplore. This particular timeline is chosen based on the following facts. (1) Since 2011, BDA has become a global phenomenon, even though the term of "Big Data" emerged back in 2007. The research before 2011 was rather insignificant in terms of volume and lacked considerable contributions to theory and practices (Manyika et al., 2011). (2) This review intends to provide the latest development in BDA-SCM research.

The initial search generated a total of 1,565 papers. After eliminating duplicated results, the total number of papers dropped to 875. Then, we checked overall relevance of the remaining papers by removing the papers that do not contain both keywords related to BDA and SCM functions in title or abstract. This screening process reduced the number of papers to 598. The remaining papers were then filtered based on inclusion and exclusion criteria. These criteria were developed and justified by the authors in order to minimize the impact of the subjective bias, as suggested by Tranfield et al. (2003). The 413 papers that met the inclusion criteria went into the final filtering with exclusion criteria.

After critically reading the introduction and discussion section of the remaining 413 papers, the following exclusion rule was applied: removing the papers that only mention the application of BDA on SCM as a fleeting point of reference or as collateral research topics. In fact, many BDA-related papers only point out the potential benefits and applications of BDA to SCM without investigating *how* they are actually implemented, i.e extracting, loading, and transferring massive datasets, and use advanced analytic techniques to support SCM decision-making. In the end, 88 papers were kept for full review. Figure 1 summarizes our systematic articles search and selection process.

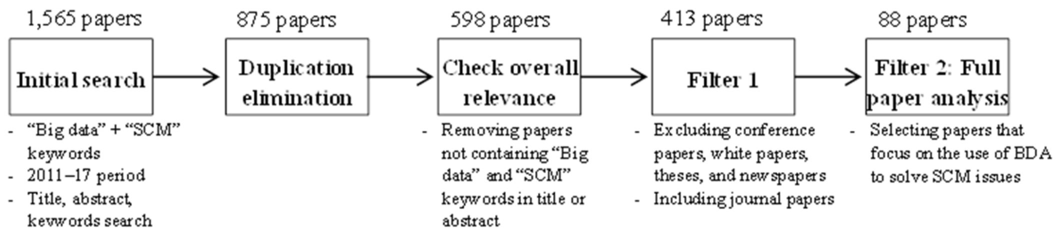


Figure 1: Systematic literature search process

2.2 Descriptive analysis

Figure 2 indicates that the number of papers published in this field have continuously increased over the last five years, and especially rocketed since 2014. This spread of publications is consistent with the frequency distribution found in Gandomi and Haider (2015). It suggests that the application of BDA on the SCM arena is a fast-growing and fruitful research field, being promoted by several special issue calls.

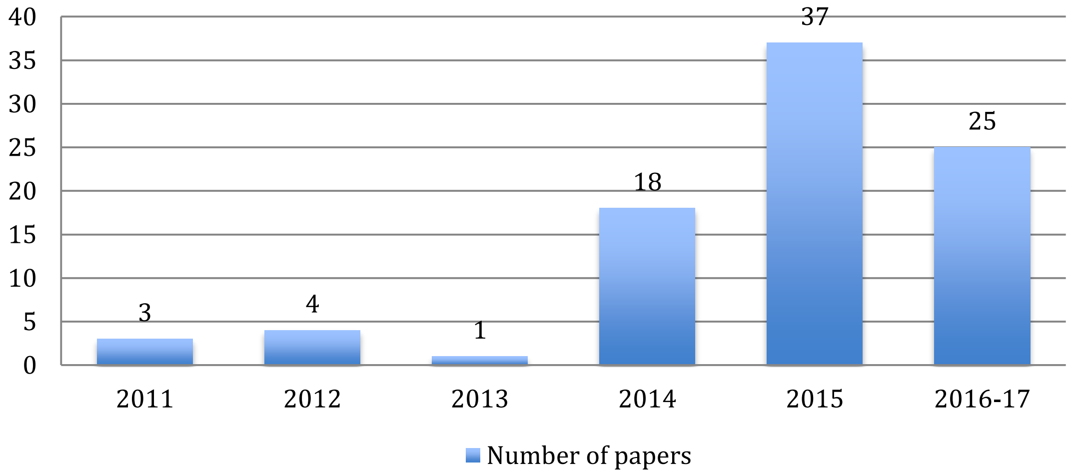


Figure 2: Distribution of reference papers by year

The selected 88 papers are from 46 different journals in which only 14 published more than one paper. Figure 3 illustrates the distribution of the reference papers in these 14 journals. It suggests that the topic is covered in a great variety of journals. Furthermore, the research topic has attracted real interest from highly regarded academics as most of these papers are published by journals with high impact factors.

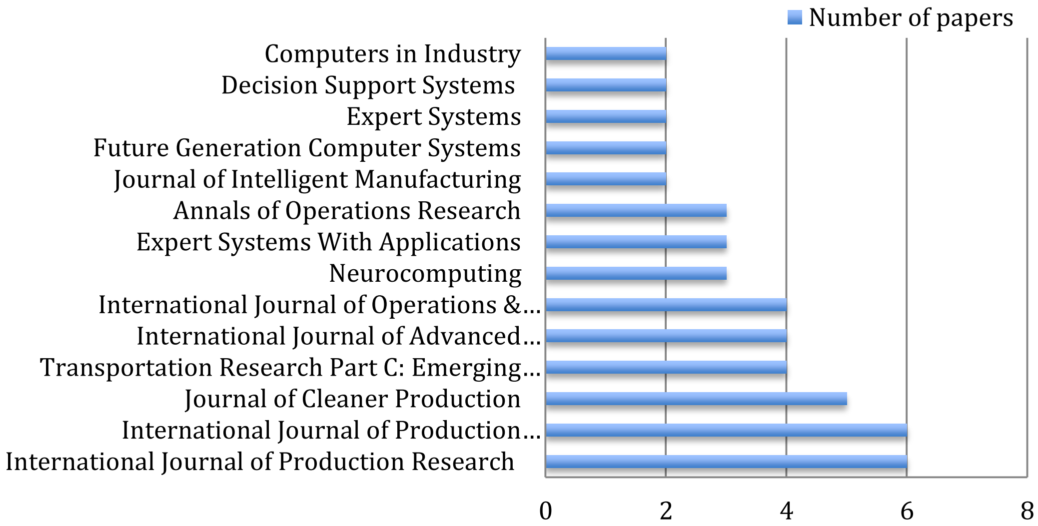


Figure 3: Distribution of reference papers by publication

2.3 Category selection

The category selection step is to conceptualize our classification framework, which is constituted by structural dimensions and analytic categories.

In order to address the proposed research questions, we select four structure dimensions with which to layer the classification framework: SCM functions, levels of analytics, BDA models, and BDA techniques. Analytic categories – i.e. key topics in each dimension – were derived deductively, based on various classification frameworks from prior literature. Table 1 presents those analytic categories.

Table 1. Literature review classification

Structure Dimension		**Analytic categories**
SC functions	Procurement	Supplier selection, sourcing cost improvement, sourcing risk management (Olson, 2015; Rozados and Tjahjono, 2014; Sanders, 2014, p.132)
	Manufacturing	Product research and development (R&D), production planning and control, quality management, maintenance and diagnosis (Meziane and Proudlove, 2000)
	Logistics/ transportation	Intelligent transportation system, logistics planning and in-transit inventory management (Wegner and Küchelhaus, 2013)
	Warehousing	Storage assignment, order picking, inventory control (Rozados and Tjahjono, 2014)
	Demand management	Demand forecasting, demand sensing, demand shaping (Chase, 2016)
Level of analytics		Descriptive, predictive, prescriptive (Saumyadipta et al., 2016, p.15)
BDA models		Visualization, association, clustering, classification, regression, forecasting, semantic analysis, optimization, simulation (Erl et al., 2016, p.181)
BDA techniques		Association rule mining, clustering algorithms, support vector machine, linear/logistics regression, neural network, fuzzy logic, Naïve Bayes, text mining, sentiment analysis, feature selection, OLAP, statistics, to name a few

In order to ensure the exhaustive categorization of each article being reviewed, there are some supplemental categories, for example, “General SCM”, “Mixed”, and “N/A”. To avoid confusion, those categories are not presented in the graphical classification framework in Figure 4.

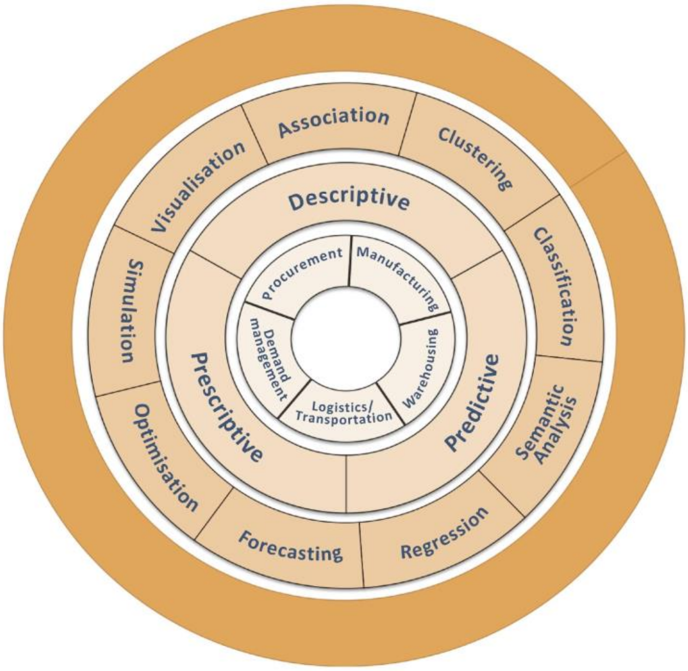


Figure 4: Classification framework

The first layer represents the key functions of SCM. In the second layer, we classify the BDA-SCM literature based on three levels of data analytics, namely descriptive, predictive, and prescriptive. This taxonomy has been widely adopted in BDA studies as it reflects complexity of both BDA-applied problems and data analytics techniques (Delen and Demirkan, 2013; Duan and Xiong, 2015). Descriptive analytics is the simplest form of BDA, describing what happened in the past, while predictive analytics predicts future events, and prescriptive analytics refers to decision-making mechanism and tools (Rehman et al., 2016). The third layer comprises the nine most common types of BDA model (Erl et al., 2016, p.181). The final layer shows the BDA techniques, which can be adopted from multiple data analytics disciplines such as data miming, machine learning, etc.

3. Material evaluation

3.1 Reviewing by SC functions

The distribution of BDA studies in each SC function is shown in Figure 5 and the classification of the examined literature by key application topics is summarized in Table 2.

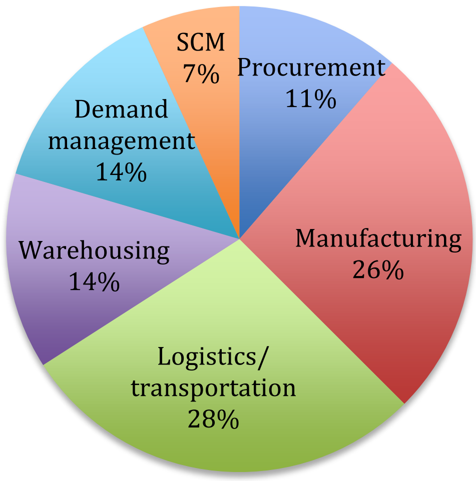


Figure 5: Distribution of each SC function in BDA

As seen in Figure 5, logistics/transportation and manufacturing have extremely dominated over the current literature on this topic, together taking up more than half of the publications. Research on the other three fundamental SC functions – warehousing, demand management and procurement – are limited, but relatively well distributed.

Among five areas of SCM, logistics/transportation (25 out of 88 papers, 28%) is the most prevalent area where BDA is used to support decision-making. The majority of research papers in this area (15 papers, 60%) focus on using BDA to develop intelligent transportation systems (ITS), while the use of BDA in supporting logistics planning has increasingly gained attention from recent research (8 papers, 32%). Only two papers (8%) are concerned with the use of BDA for inventory management during the in-transit logistics process.

Another area in which BDA applications have been studied extensively is manufacturing. It takes up 26% of total publications. Out of the 23 manufacturing-related articles, more than half (13 papers, 57%) are related to production planning and control, while production R&D, maintenance and diagnosis each represents 26% of publications (6 papers each). It is worth noting that the use of BDA for quality control during the manufacturing process is little discussed, appearing in only 4 papers.

In demand management literature, sensing current demand (7 papers) and shaping future demand (6 papers) are among the most prominent initiatives of BDA, while, surprisingly, demand forecasting is seldom the focus of the study (only 3 papers).

In warehousing, BDA has been widely recognized to improve storage assignment (5 papers) and inventory management (5 papers), whilst the use of BDA to support the order-picking process is under-examined (3 papers).

BDA research focusing on procurement is well balanced over three major issues: supplier selection (5 papers), sourcing improvement (4 papers), and sourcing risk management (4 papers).

While most studies in the literature examine the application of BDA to specific SC functions, we found 6 papers (7% of total publications) that analyse BDA applications while considering SCs as multi-level interconnected networks. These papers address different SC issues concerning resilience, sustainability, risk management, and agility.

Table 2: Summary of literature by SC function

SC function	Key activity	Paper
Procurement	Supplier selection	Choi et al., 2016; Huang and Handfield, 2015; Jain et al., 2014; Kuo et al., 2015; Mori et al., 2012
	Sourcing cost improvement	Ahiaga-Dagbui and Smith, 2014; Huang and Handfield, 2015; Kuo et al., 2015; Tan and Lee, 2015
	Sourcing risk management	Ghedini Ralha and Sarmento Silva, 2012; Huang and Handfield, 2015; Ling Ho and Wen Shih, 2014; Miroslav et al., 2014
Manufacturing	Product R&D	Bae and Kim, 2011; Do, 2014; Lei and Moon, 2015; Opresnik and Taisch, 2015; Tan et al., 2015; Zhang et al., 2017
	Production planning & control	Chien et al., 2014; Krumeich et al., 2016; Li et al., 2016; Wang and Zhang, 2016; Zhang et al., 2017; Zhong et al., 2015; Shu et al., 2016; Dai et al., 2012; Zhang et al., 2015; Zhong et al., 2015; Zhong et al., 2016; Zhong et al., 2015; Helo and Hao, 2017
	Quality management	Krumeich et al., 2016; Wang et al., 2016; Zhang et al., 2017; Zhang et al., 2015
	Maintenance & diagnosis	Shu et al., 2016; Guo et al., 2016; Kumar et al., 2016; Zhang et al., 2017; Wang et al., 2016; Wang et al., 2015
Warehousing	Storage assignment	Chuang et al., 2014; Li, Moghaddam, et al., 2016; Tsai and Huang, 2015; Chiang et al., 2011; Chiang et al., 2014
	Order picking	Ballestín et al., 2013; Chuang et al., 2014; Li et al., 2016
	Inventory control	Alyahya et al., 2016; Hofmann, 2015; Hsu et al., 2015; Huang and Van Mieghem, 2014; Lee et al., 2016; Stefanovic, 2015
Logistics and transportation	ITS	Cui et al., 2016; Li et al., 2015; Shi and Abdel-Aty, 2015; St-Aubin et al., 2015; Toole et al., 2015; Wang et al., 2016; Xia et al., 2016; Yu and Abdel-Aty, 2014; Zangenehpour et al., 2015; Dobre and Xhafa, 2014; Ehmke et al., 2016; Sivamani et al., 2014; Toole et al., 2015; Zhang et al., 2016; Hsu et al., 2015
	Logistics planning	Lee, 2016; Prasad et al., 2016; Yan-Qiu and Hao, 2016; Zhao et al., 2016; Shan and Zhu, 2015; Tu et al., 2015; Li et al., 2014; Mehmood et al., 2017
	In-transit inventory management	Ting et al., 2014; Delen et al., 2011
Demand management	Demand forecasting	Berengueres and Efimov, 2014; Chong et al., 2016; Jun et al., 2014; Li, Ch'ng, et al., 2016; Ma et al., 2014
	Demand sensing	Berengueres and Efimov, 2014; Chong et al., 2016; Fang and Zhan, 2015; He et al., 2015; Li, Ch'ng, et al., 2016; Salehan and Kim, 2016; Wang et al., 2014

	Demand shaping	Chong et al., 2016; He et al., 2015; Marine-Roig and Anton Clavé, 2015; Salehan and Kim, 2016; Schmidt et al., 2014
General SCM	Ong et al., 2015; Papadopoulos et al., 2017; Sheffi, 2015; Ting et al., 2014; Wu et al., 2017; Zhao et al., 2016, 2015	

3.2 Reviewing by level of analytics

The rationale of using this taxonomy is to examine the extent to which BDA is being used to support decision-making processes, as well as understanding what types of SC problems are being solved.

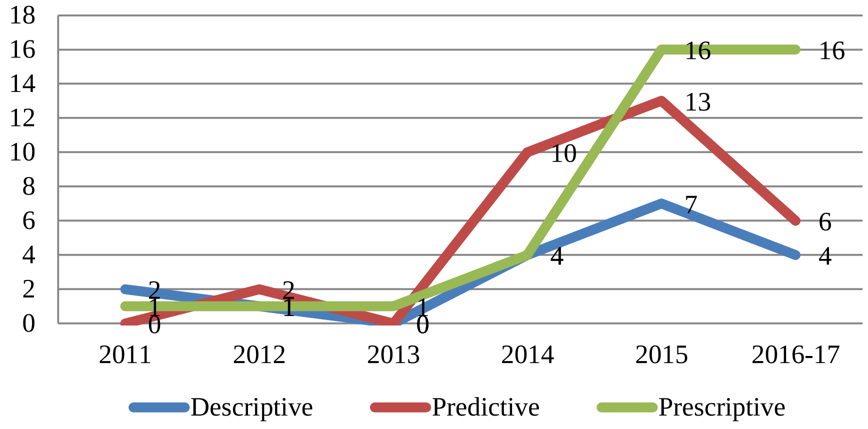


Figure 6: Distribution of analytics level by year

Figure 6 depicts the popularity of each analytics type by year. Although the trend from 2011 to 2013 is underrepresented due to the insufficient number of BDA-SCM studies, we can still see that the majority of studies in this early stage used BDA for descriptive analytics, while predictive and prescriptive analytics were little discussed. However, these minorities have changed drastically along with the upsurge in the publication level since 2013. Particularly, predictive analytics dominated in 2014, accounting for 55% of publications (10 out of 18 papers). Meanwhile, prescriptive analytics has been the fastest growing since 2014 and has become the most common type in 2015 with 16 out of 36 papers (44.4%). Descriptive analytics has also aligned with these upward trends, but not remarkably soared like prescriptive analytics.

Table 3: Level of analytics in each SC function

	Descriptive	Predictive	Prescriptive
Procurement	3	5	2
Manufacturing	4	7	12
Warehousing	3	3	6
Logistics/transportation	4	6	15
Demand management	2	10	0
Other SC topics	2	0	4
Total papers	18	31	39

To reveal more insights from this taxonomy of BDA, we further investigate how each level of analytics has been studied in each specific SC domain. Table 3 presents the result. Overall, prescriptive analytics is the most discussed type in the examined literature, taking up 44.3% of publications (39 out of 88 papers), while predictive analytics is just behind with 31 papers (35.2%), with descriptive analytics taking up the least coverage (18 papers, 20.5%). In particular, the result found that manufacturing (12 papers) and logistics/transportation (15 papers) are those areas which have mainly contributed to the prominence of prescriptive analytics, thanks to the increasing adoption of various state-of-the-art systems such as Cyber Physical System (CPS) and ITS. Meanwhile, predictive analytics is the most often common type used in demand management (6 out of 12 papers) and procurement (4 out of 10 papers). Accurate demand forecasting and early detection of various sourcing risks are among the foremost applications of BDA-enabled predictive models in these two areas. It should be noted that prescriptive applications of BDA in demand management and procurement are seldom studied.

3.3 Literature review by type of BDA model

The result of the literature review based on BDA models is summarized in Table 4.

Table 4: Distribution of articles by BDA models

	Procurement	Manufacturing	Warehousing	Logistics/ transportation	Demand management	General SCM	Total papers
Optimization	1	5	4	11	0	0	**21**
Mixed/others	0	4	0	3	0	4	**11**
Classification	2	5	0	3	0	0	**10**
Association	2	1	3	2	2	0	**10**
Semantic analysis	1	1	0	0	6	1	**9**
Forecasting	1	0	2	2	3	0	**8**
Simulation	1	3	2	1	0	0	**7**
Clustering	1	0	0	2	1	1	**5**
Regression	0	1	1	1	0	0	**3**
Visualization	0	2	0	0	0	0	**2**
N/A	1	1	0	0	0	0	**2**

As discussed in Section 3.2, since the majority of reviewed papers focus on the prescriptive applications of BDA, the adoption of optimization and simulation modelling to support decision-making seems to be natural. Optimization is adopted in 21 papers (Table 4), but surprisingly enough, simulation is only applied in 7 papers. Route optimization and logistics planning in logistics/transportation areas (11 papers) have mainly contributed to the dominance of optimization, whereas in contrast, there is no paper found in the examined literature that uses simulation in this area.

For the predictive level of analytics, classification is the most used BDA model in SCM context (10 papers). The model aims to classify a huge set of data objects into predefined categories, thereby generating predictions with high levels of accuracy (Mastrogiannis et al., 2009). Classification has been largely employed in manufacturing (5 out of 10 papers, or 50%), logistics/transportation (3 papers, 30%), and procurement (2 papers, 20%). Other popular models for predictive analytics are semantic analysis (9 papers) and forecasting (8 papers). While the application scope of forecasting models is quite diversified – ranging from demand management, warehousing, and logistics/transportation to procurement – more than half of studies using semantic analysis (6 out of 9 papers, or 66.7%) are still limited to demand management area.

At the descriptive level of BDA, association is the most used approach (10 out of 88 papers, or 11.36%), which refers to the discovery of recurring and strong relationships among items in large-scale datasets. It should be noted that association is also the most versatile model in the BDA-SCM field as it can support decision-making in every stage of the SC process, from procurement, manufacturing, warehousing, logistics/transportation to demand management. Visualization, surprisingly, is the least used model in descriptive analytics as well as in the BDA-SCM literature as a whole. There are only two papers in the examined literature studying this type of model as the main research focus.

Finally, a number of papers are classified into mixed/other models (11 papers, 12.5% of publications). Those papers fall into three SC functions, namely manufacturing (4 papers), logistics/transportation (3 papers), and general SCM (4 papers), as seen in Table 4.

3.4 Review of articles by BDA techniques

This section focuses on the types of BDA techniques and algorithms that have been used to develop the BDA models discussed in Section 3.3. The result is summarized in Table 5.

Table 5: Distribution of articles by BDA techniques

	Association	Clustering	Classification	Regression	Forecasting	Semantic analysis	Optimization	Simulation	Visualization	Mixed	Total papers
Association rule mining	8		1				4				**13**
Mixed*		1	2				3			7	**13**
Heuristic approach							7	1			**8**
Spatial-temporal visualization							5	1	1		**7**
Decision tree	1		3		3						**7**
Support vector machine			6								**6**
Sentiment analysis						3			2		**5**
Logistic regression			1	3	1						**5**
Generic Algorithm							4				**4**
Neural network					3	1					**4**
Text mining		1	1			2					**4**
K-mean clustering		1	1		1			1			**4**
Time-series forecasting					3		1				**4**

Table 5 highlights the prevalent use of some techniques with particular types of BDA models. For instance, support vector machine (SVM) (6 papers) prevails in classification models, while heuristic approaches (7 papers) along with spatial/temporal-based visual analysis (5 papers) are key methods to develop BDA-driven optimization models. In addition, there are some techniques that can be used flexibly with different types of BDA modelling. For example, the K-means clustering algorithm is among the most adaptable techniques as it can be applied in clustering, classification, forecasting, and simulation models. Another versatile technique, and also the most frequently used technique in the BDA-applied SCM literature, is association rule mining (ARM). This method has been extensively studied in descriptive association models (7 papers) but is increasingly used to facilitate more complicated analytics models in predictive and prescriptive levels such as classification (1 paper) and optimization (4 papers).

4. Results and discussion

(1) In what areas of SCM is BDA being applied?

In logistics, transportation management prevails, with particular focus on three fundamental functions of ITS: routing optimization, real-time traffic operation monitoring, and proactive safety management. It is noteworthy that the BDA-driven routing problem is mainly studied in static environments based on historical databases (Ehmke et al., 2016; Zhang et al., 2016), while the use of BDA for dynamic routing optimization in real-time contexts is only conceptualized in some theoretical platform-based papers such as Sivamani et al. (2014) and Hsu et al. (2015). Moreover, the application of BDA on logistics network planning has recently gained more attention, but is still under-examined in both strategic and operational levels (Zhao et al., 2016; Mehmood et al., 2017). Finally, the monitoring and controlling of product conditions through sensors during the in-transit process is seldom addressed (Delen et al., 2011; Ting et al., 2014).

Production planning and control is currently receiving the most research interest, and the application of BDA theories and tools on this topic is at a relatively mature stage (Wang and Zhang, 2016; Zhong et al., 2015). Although BDA adoption in product R&D and equipment diagnosis and maintenance is less often studied, papers in this area make a significant contribution to predictive and prescriptive analytics in manufacturing research (Lei and Moon, 2015; Wang et al., 2015; Wang and Zhang, 2016; Zhang et al., 2017). Moreover, research on BDA-enabled quality control during manufacturing processes is rather limited (Krumeich et al., 2016; Zhang et al., 2015).

With regards to the warehousing operation, it is well documented in the literature that storage assignment and inventory control are taking advantage of BDA. However, inventory control-related dynamics, such as the Bullwhip effect, have just recently been discussed in theory (Hofmann, 2015). Furthermore, only few studies address order-picking problems in BDA-enabled warehousing (Ballestín et al., 2013; Chuang et al., 2014). The study of how BDA can optimize order-picking processes, such as order batching, routing, and sorting, is still scarce.

Studies of BDA in the procurement area are evenly spread over the three major applications of supplier selection, sourcing cost improvement, and sourcing risk analysis. BDA has been widely adopted to facilitate the supplier selection process and recent efforts have been made to integrate this activity with order allocation problems and to reduce sourcing costs (Kuo et al., 2015). In terms of sourcing risk management, most studies have only exploited the benefit of BDA in accurately detecting procurement risk based on a sizable supplier database, while models and decision support systems (DSS) that provide proactive prevention actions are still lacking (Ghedini Ralha and Sarmento Silva, 2012; Miroslav et al., 2014).

The examined literature provides numerous contributions in terms of capturing demand changes in real-time. BDA can help in sensing demand behaviours to increase the agility and accuracy of demand forecasting (Fang and Zhan, 2015; Salehan and Kim, 2016; Wang et al., 2014). Another common application of BDA in demand management is shaping demand to be aligned with production and logistics capacity. However, current studies on this issue have taken a marketing intelligence perspective rather than an operational SCM perspective (Marine-Roig and Anton Clavé, 2015; Schmidt et al., 2015).

Finally, the review shows that recent research has increasingly recognized the importance of studying BDA with a holistic perspective mindful of SCM as a multi-level, inextricably interlinked system. Most of those studies examined the SC integration in the context of SC resilience (Papadopoulos et al., 2017; Sheffi, 2015), sustainability (Papadopoulos et al., 2017; Wu et al., 2017), risk management (Ong et al., 2015), and agility (Giannakis and Louis, 2016). However, the research on this issue still strongly emphasizes theoretical development with limited studies of advanced data-mining modelling.

(2) At what level of analytics is BDA used in these SCM areas?

The rationale of this research question is to investigate the level of data analytics required in the SC application, as well as indicating the types of problem being solved.

In the trend analysis, the results show that prescriptive analytics is the most common and fastest growing area in the BDA-driven SCM, which is closely followed by predictive analytics, while descriptive analytics is receiving less consideration. To be more specific, logistics/transportation, manufacturing, and warehousing domains are the major contributors of prescriptive analytics, thanks to the increasing adoption of various state-of-the-art systems such as CPS in Industry 4.0 (Wang et al., 2016; Helo and Hao, 2017) and ITS (Wang et al., 2015). On the other hand, predictive analytics is still the primary actor in demand management and procurement, especially for demand forecasting and sourcing risk management, while prescriptive analytics is still rarely discussed (Ghedini Ralha and Sarmento Silva, 2012).

(3) What types of BDA models are used in SCM?

Optimization is the most popular approach when it comes to prescriptive analytics, especially in the logistics and transportation area. As aforementioned, literature in logistics/transportation has made little insight so far on real-time routing optimization based on streamline data. On the other hand, the study of real-time optimization appears to be quite mature in the manufacturing domain with the use of modelling and simulation to develop a real-time production control system, based on streamline context-aware data, generated from tracking devices such as RFID (Babiceanu and Seker, 2016; Kumar et al., 2016). It is highly possible for transportation controllers and warehouse operators to adapt a similar approach of modelling and simulation to optimize routing problem in real-time, as suggested in (Wang et al., 2016).

Classification is the most common approach at the predictive analytics level and has been widely applied in manufacturing to support production planning and control (Chien et al., 2014; Wang and Zhang, 2016) and equipment maintenance and diagnosis (Kumar et al., 2016; Shu et al., 2016; Wang et al., 2016). This type of BDA model also plays a key role in logistics/transportation (Li et al., 2014; Yu and Abdel-Aty, 2014; Zangenehpour et al., 2015) and procurement research (Ling Ho and Wen Shih, 2014; Mori et al., 2012) but apparently, current studies in those areas have not fully exploited the advantages of classification.

Another popular model for predictive analytics is semantic analysis, but its scope of application is still considerably limited to demand sensing. It could be beneficial for future research to extend the application of this approach to more SC and operational management such as fraud detection (Miroslav et al., 2014) and behaviour-based safety analysis (Guo et al., 2016).

For descriptive analytics, association is the most widespread as it has been applied throughout every stage of the SC process, from procurement (Ghedini Ralha and Sarmento Silva, 2012; Jain et al., 2014), manufacturing (Bae and Kim, 2011), warehousing (Chiang et al., 2011, 2014; Chuang et al., 2014), and logistics/transportation (Cui et al., 2016), to demand management (Jin et al., 2016). Notably, the visualization model is rarely considered as the main focus of a study (Zhong et al., 2016; Zhong et al., 2015), but is commonly used as a complement to other advanced data-mining models (Shan and Zhu, 2015; Zhang et al., 2016).

Finally, the review classified 10 papers (11.6% of total papers) under “mixed/others” models. Most of those papers focus on the intelligent DSS that enables real-time control of the entire operational process in manufacturing (Dai et al., 2012; Krumeich et al., 2016; Zhang et al., 2017; Zhong et al., 2016), logistics/transportation (Delen et al., 2011; Dobre and Xhafa, 2014; Hsu, Lin, et al., 2015), and SC agility (Ong et al., 2015; Papadopoulos et al., 2017). Not surprisingly, mathematical models are missing in many of those papers since such state-of-the-art systems normally require a complex mixture of various models and techniques from a number of data-mining disciplines.

(4) What BDA techniques are employed to develop these models?

There is a wide range of BDA techniques and algorithms that have been used in the SCM context. Some of them are prevalent in particular modelling approaches, for example, SVM in classification models, heuristics approach in optimization models, and neural networking in forecasting models. The review also identifies some versatile techniques that can be adapted to different types of models. For instance, K-means clustering algorithm is among the most adaptable techniques as it can be adopted in clustering (St-Aubin et al., 2015; Tan and Lee, 2015), classification (Chien et al., 2014), forecasting (Stefanovic, 2015), and modelling and simulation (Lei and Moon, 2015). In those studies, K-means is often performed in the initial phase of the data analytics process to partition the raw heterogeneous datasets into more homogenous segments. Studies find that advanced data-mining techniques such as decision trees and neural network would develop more accurate predictive models by leveraging the result of cluster analysis (Krumeich et al., 2016; Lei and Moon, 2015; Stefanovic, 2015).

Another highly adaptable BDA technique is ARM. Apart from descriptive analytics, scholars have increasingly used this method to facilitate more complicated analytics at the predictive and prescriptive level. As yet, we have only found one paper, Ling Ho and Wen Shih (2014), that conceptualizes the notion of using ARM along with the decision tree technique to develop the highly accurate prediction models for procurement risk. However, the mathematical model and algorithm to put this idea into practice is still missing. Interestingly, while the predictive application of ARM is little discussed in the literature, its contribution to prescriptive analytics seems to be more recognized. Indeed, we found 4 papers incorporating ARM with optimization models to effectively solve allocation problems in various SC areas. For example, Li et al. (2016) use ARM and Generic algorithm (GA) to optimize storage assignments, thus enhancing order-picking processes. Tsai and Huang (2015) optimize shelf space allocation by using ARM, sequential pattern mining, and the combinatorial optimization approach. For logistic and transportation planning, Lee (2016) uses ARM to extract purchase patterns and perform if-then-else rules to predict customer purchasing behaviour, thus proposing the GA approach to optimize anticipatory shipping assignment. Finally, ARM can also be used in the hybrid optimization problems of supplier selection and order allocation (Kuo et al., 2015).

Not surprisingly, there are a large number of papers under the "mixed/others" category (i.e the combination of more than three different methods) since there is no single technique that is fully capable of managing the complex and diverse nature of BD (Chen and Zhang, 2014).

5. Future direction

The findings discussed above suggest some future directions to capitalize the research development of BDA applications in the SCM context.

1. Further investigation of BDA application to SC function level

The review suggests a number of research gaps in each SC function. For example, quality control in manufacturing, dynamic vehicle routing and in-transit inventory management in logistics/transportation, order picking and inventory control systems in warehousing, demand shaping in SC and operational research, and procurement are some of those areas that are currently much less discussed.

2. Functional alignment strategy for the horizontal integration of BDA-driven SC

SCM is the multi-level process by which all functions are interlinked. Hence, fragmental efforts of BDA adoption to only one or two functions will not yield any significant and long-lasting competitive advantage. To avoid such fragmented efforts, the entire SC should be horizontally integrated by aligning BDA applications in different functions effectively. For example, production and logistics planning could incorporate with real-time demand sensing for cost reduction and higher service level.

Indeed, alignment dissolves the boundary across functions.

To facilitate the horizontal integration throughout the SC, future BDA research should focus more on cross-functional problems such as vehicle routing and facility location, supplier selection and order allocation, demand-driven storage assignment and order picking.

3. Three levels of analytics should be equally examined

As aforementioned, current research focuses more on prescriptive analytics than descriptive and predictive analytics. Nonetheless, the application of BDA in any subject, not just SCM, is always a linear process. In this process, the performance of prescriptive analytics would heavily rely on those of descriptive and predictive analytics as they dictate the value of critical parameters in prescriptive models (Duan and Xiong, 2015). To catalyse the rapid progression of BDA application in SCM, future research should balance the focus across all three levels of analytics.

4. Combining different data analytic techniques to develop more advanced and adaptive BDA models for DSS

The literature review has identified a number of BDA models commonly used in SCM applications as well as popular and versatile BDA techniques to build those models. Dynamic optimization and simulation modelling should be further investigated in the context of BDA as they are baseline approaches for prescriptive analytics and DSS.

Moreover, although the literature has extensively adopted visualization techniques as supplement techniques to predictive and prescriptive models, little attention has been paid on improving data visualization techniques. Future research should call for this gap because visualizing complex BD would expedite decision making.

5. Application of BDA on closed-loop SCM

It is rather surprising that research on applying BDA on reverse logistics and the closed-loop supply chain (CLSC) is scarce. This might be due to the fact that collecting data for used products is extremely hard, which hinders introducing BDA into CLSC management. Nevertheless, the development of new technologies such as IoT, machine-to-machine, would be able to overcome this barrier. The BDA that has already been applied in product lifecycle design and assessment (Ma et al., 2014; Song et al., 2016) would be useful for predicting product returns and estimating the return quality. This is important for capacity planning and remanufacturing scheduling in a reverse logistics system.

Managing a CLSC has always been challenging due to the uncertainties and possible conflicting goals, i.e. profit vs. environment vs. social wellbeing. In this sense, BD would be useful in understanding people perception, devising multi-KPIs, monitoring operational process, and then taking corresponding actions. To achieve this, developing the knowledge database, tools, and techniques of BDA must be on the research agenda.

6. BD-driven business models in SC

BD revolutionizes SC business models. On the one hand, it shortens the SC layers. On the other hand, it expands revenue streams from existing products to servitization, and creates new revenue streams from entirely new (data) products (Opresnik and Taisch, 2015). Nevertheless, the ecosystem that supports new business models is underdeveloped; the enablers and obstacles of new business models remain unclear.

This leads to the research questions that we propose for future study: (1) what are the BD strategies in SCM? (2) How to increase the VALUE of BD, the most important

‘V’; (3) how various stakeholders contribute to adding value of BD and what is the revenue sharing mechanism among the stakeholders in SC; (4) what is the dynamic impact of new business model on SC performance as whole; (5) what are the tipping points that transfer a conventional business model to a BD-driven business model.

7. New tools and BDA techniques for distributed SC and distributed computation

Cloud computing, countless sensors around us, distributed service resources, and distributed operational processes generate voluminous amounts of data. Coordinating a *distributed* SC and managing the complex procedures of different BDA are challenging (Li et al., 2016).

The review suggests that the majority of current research has been focusing on the one-location, one-computer scenario. This doesn’t reflect the reality of *distributed* systems. We call for research on developing SC system-wide feedback and coordination based on BDA to optimize system performance (Wang et al., 2016). It is anticipated that the framework and operational mechanism of today’s smart factory would be scalable to the entire SC. This SC system should be a self-organized reconfiguration with BD-based feedback and coordination without, or with very limited, human intervention. To achieve this, besides hardware and infrastructure, future research should develop more efficient data-intensive techniques and technologies (Chen and Zhang, 2014).

6. Conclusion

Based on the content analysis methodology of Mayring (2008), this literature review examined 88 journal papers to provide a full picture of where and how BDA has been applied within the SCM context. In particular, we developed a classification framework based on four research questions: (1) in what areas of SCM is BDA being applied? (2) At what level of analytics is BDA used in these SCM areas? (3) What types of BDA models are used in SCM? And, finally, (4) what BDA techniques are employed to develop these models? Addressing these questions, the discussion has highlighted a number of research gaps and future directions for BDA applications to catalyse the research development of the topic.

One of the limitations of this paper is that the categorization in the classification framework remains interpretative, as this could lead to concern on subjective bias. This is also one of the well-established issues of the content analysis method despite a number of validations being carried out (Seuring, 2013). Another concern could be on the time period selected in this review, from 2011 to 2016. Although Figure 1 and Figure 2 indicate the fact that BDA literature has made major contributions since 2013, the topic of data-driven SCM and data analytics have been studied long over a decade.

References

Addo-Tenkorang, R. and Helo, P.T. (2016), “Big data applications in operations/supply-chain management: A literature review”, *Computers & Industrial Engineering*, Elsevier Ltd, Vol. 101, pp. 528–543.

Ahiaga-Dagbui, D.D. and Smith, S.D. (2014), “Dealing with construction cost overruns using data mining”, *Construction Management and Economics*, Vol. 32 No. 7–8, pp. 682–694.

Alyahya, S., Wang, Q. and Bennett, N. (2016), “Application and integration of an RFID-enabled warehousing management system – a feasibility study”, *Journal of Industrial Information Integration*, Elsevier Inc., Vol. 4, pp. 15-25, doi:10.1016/j.jii.2016.08.001.

Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M. a. S. and Buyya, R. (2015), “Big Data computing and clouds: Trends and future directions”, *Journal of Parallel and Distributed Computing*, Elsevier Inc., Vol. 79–80, pp. 3–15.

Babiceanu, R.F. and Seker, R. (2016), “Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook”, *Computers in Industry*, Elsevier B.V., Vol. 81, pp. 128–137.

Bae, J.K. and Kim, J. (2011), “Product development with data mining techniques: A case on design of digital camera”, *Expert Systems with Applications*, Elsevier Ltd, Vol. 38 No. 8, pp. 9274–9280.

Ballestín, F., Pérez, Á., Lino, P., Quintanilla, S. and Valls, V. (2013), “Static and dynamic policies with RFID for the scheduling of retrieval and storage warehouse operations”, *Computers & Industrial Engineering*, Vol. 66 No. 4, pp. 696–709.

Berengueres, J. and Efimov, D. (2014), “Airline new customer tier level forecasting for real-time resource allocation of a miles program”, *Journal Of Big Data*, Vol. 1 No. 1, p. 3.

Chase, C.W. (2016), *Next Generation Demand Management: People, Process, Analytics, and Technology*, Wiley Online Library, available at: 4.

https://books.google.co.uk/books?id=CkbhCgAAQBAJ&printsec=frontcover#v=onepage&q=demand shaping&f=false.

Chen, C.L.P. and Zhang, C.Y. (2014a), “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data”, *Information Sciences*, Elsevier Inc., Vol. 275, pp. 314–347.

Chen, C.L.P. and Zhang, C.-Y. (2014b), “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data”, *Information Sciences*, Elsevier Inc., Vol. 275, pp. 314–347.

Chiang, D.M.-H., Lin, C.-P. and Chen, M.-C. (2011), “The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres”, *Enterprise Information Systems*, Vol. 5 No. 2, pp. 219–234.

Chiang, D.M.-H., Lin, C.-P. and Chen, M.-C. (2014), “Data mining based storage assignment heuristics for travel distance reduction”, *Expert Systems*, Vol. 31 No. 1, pp. 81–90.

Chien, C., Diaz, A.C. and Lan, Y. (2014), “A data mining approach for analyzing semiconductor MES and FDC data to enhance overall usage effectiveness (OUE)”, *International Journal of Computational Intelligence Systems*, Vol. 7 No. sup2, pp. 52–65.

Choi, Y., Lee, H. and Irani, Z. (2016), “Big data-driven fuzzy cognitive map for prioritising IT service procurement in the public sector”, *Annals of Operations Research*, Springer US, Vol. 243 No. 1–2, pp. 1–30.

Chong, A.Y.L., Li, B., Ngai, E.W.T., Ch’ng, E. and Lee, F. (2016), “Predicting online product sales via online reviews, sentiments, and promotion strategies”, *International Journal of Operations & Production Management*, Vol. 36 No. 4, pp. 358–383.

Chuang, Y.F., Chia, S.H. and Wong, J.Y. (2014), “Enhancing order-picking efficiency through data mining and assignment approaches”, *WSEAS Transactions on Business and*

Economics, Vol. 11 No. 1, pp. 52–64.

Cui, J., Liu, F., Hu, J., Janssens, D., Wets, G. and Cools, M. (2016), "Identifying mismatch between urban travel demand and transport network services using GPS data: A case study in the fast growing Chinese city of Harbin", *Neurocomputing*, Elsevier, Vol. 181, pp. 4–18.

Dai, Q., Zhong, R., Huang, G.Q., Qu, T., Zhang, T. and Luo, T.Y. (2012), "Radio frequency identification-enabled real-time manufacturing execution system: a case study in an automotive part manufacturer", *International Journal of Computer Integrated Manufacturing*, Vol. 25 No. 1, pp. 51–65.

Delen, D. and Demirkan, H. (2013), "Data, information and analytics as services", *Decision Support Systems*, Elsevier B.V., Vol. 55 No. 1, pp. 359–363.

Delen, D., Erraguntla, M., Mayer, R.J. and Wu, C.-N. (2011), "Better management of blood supply-chain with GIS-based analytics", *Annals of Operations Research*, Vol. 185 No. 1, pp. 181–193.

Do, N. (2014), "Application of OLAP to a PDM database for interactive performance evaluation of in-progress product development", *Computers in Industry*, Elsevier B.V., Vol. 65 No. 4, pp. 636–645.

Dobre, C. and Xhafa, F. (2014), "Intelligent services for Big Data science", *Future Generation Computer Systems*, Elsevier B.V., Vol. 37, pp. 267–281.

Duan, L. and Xiong, Y. (2015), "Big data analytics and business analytics", *Journal of Management Analytics*, Taylor & Francis, Vol. 2 No. 1, pp. 1–21.

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*, Vol. 84, No. 1–4, pp. 631–645.

Dutta, D. and Bose, I. (2015), "Managing a Big Data project: The case of Ramco Cements

Limited”, *International Journal of Production Economics*, Elsevier, Vol. 165, pp. 293–306.

Ehmke, J.F., Campbell, A.M. and Thomas, B.W. (2016), “Data-driven approaches for emissions-minimized paths in urban areas”, *Computers & Operations Research*, Elsevier, Vol. 67, pp. 34–47.

Emani, C.K., Cullot, N. and Nicolle, C. (2015), “Understandable Big Data: A survey”, *Computer Science Review*, Elsevier Inc., Vol. 17, pp. 70–81.

Erl, T., Khattak, W. and Buhler, P. (2016), *Big Data Fundamentals*, PRENTICE HALL, available at: file:///C:/2016-2017/Research/Paper/COR_Jimmy and Vir/References/Referencing for Li/37.pdf.

Fang, X. and Zhan, J. (2015), “Sentiment analysis using product review data”, *Journal of Big Data*, Journal of Big Data, Vol. 2 No. 1, p. 5.

Gandomi, A. and Haider, M. (2015), “Beyond the hype: Big data concepts, methods, and analytics”, *International Journal of Information Management*, Elsevier Ltd, Vol. 35 No. 2, pp. 137–144.

Ghedini Ralha, C. and Sarmento Silva, C.V. (2012), “A multi-agent data mining system for cartel detection in Brazilian government procurement”, *Expert Systems with Applications*, Vol. 39 No. 14, pp. 11642–11656.

Giannakis, M. and Louis, M. (2016), “A multi-agent based system with big data processing for enhanced supply chain agility”, *Journal of Enterprise Information Management*, Vol. 29 No. 5, pp. 706–727.

Govindan, K., Soleimani, H. and Kannan, D. (2014), “Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future”, *European Journal of Operational Research*, Vol. 240 No. 3, pp. 603–626.

Gunasekaran, A., Kumar Tiwari, M., Dubey, R., & Fosso Wamba, S. (2016), "Big data and predictive analytics applications in supply chain management", *Computers and Industrial Engineering*, 101(C), 525–527.

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.

Guo, S.Y., Ding, L.Y., Luo, H.B. and Jiang, X.Y. (2016), "A Big-Data-based platform of workers' behavior: Observations from the field", *Accident Analysis & Prevention*, Elsevier Ltd, Vol. 93, pp. 299–309.

Hazen, B.T., Skipper, J.B., Boone, C.A. and 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*, Springer US, Vol. May, pp. 1–11.

He, W., Wu, H., Yan, G., Akula, V. and Shen, J. (2015), "A novel social media competitive analytics framework with sentiment benchmarks", *Information & Management*, Vol. 52 No. 7, pp. 801–812.

Helo, P. and Hao, Y. (2017), "Cloud manufacturing system for sheet metal processing", *Production Planning & Control*, Vol. 28 No. May, pp. 524–537.

Hofmann, E. (2015), "Big data and supply chain decisions: the impact of volume, variety and velocity properties on the bullwhip effect", *International Journal of Production Research*, Vol. 7543 No. December 2015, pp. 1–19.

Hsu, C.-Y., Lin, S.-C. and Chien, C.-F. (2015), "A back-propagation neural network with a distributed lag model for semiconductor vendor-managed inventory", *Journal of Industrial and Production Engineering*, Vol. 32 No. 3, pp. 149–161.

Hsu, C.-Y., Yang, C.-S., Yu, L.-C., Lin, C.-F., Yao, H.-H., Chen, D.-Y., Robert Lai, K., et al. (2015), "Development of a cloud-based service framework for energy conservation in a sustainable intelligent transportation system", *International Journal of Production*

Economics, Elsevier, Vol. 164, pp. 454–461.

Huang, T., Lan, L., Fang, X., An, P., Min, J. and Wang, F. (2015), "Promises and Challenges of Big Data Computing in Health Sciences", *Big Data Research*, Elsevier Inc., Vol. 2 No. 1, pp. 2–11.

Huang, T. and Van Mieghem, J.A. (2014), "Clickstream Data and Inventory Management: Model and Empirical Analysis", *Production and Operations Management*, Vol. 23 No. 3, pp. 333–347.

Huang, Y.-Y. and Handfield, R.B. (2015), "Measuring the benefits of ERP on supply management maturity model: a 'big data' method", *International Journal of Operations & Production Management*, Vol. 35 No. 1, pp. 2–25.

IBM. (2012), "What is Big Data?.", *IBM Corporation*.

IDC. (2012), *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East.*

Jain, R., Singh, A.R., Yadav, H.C. and Mishra, P.K. (2014), "Using data mining synergies for evaluating criteria at pre-qualification stage of supplier selection", *Journal of Intelligent Manufacturing*, Vol. 25 No. 1, pp. 165–175.

Jin, J., Liu, Y., Ji, P. and Liu, H. (2016), "Understanding big consumer opinion data for market-driven product design", *International Journal of Production Research*, Vol. 54 No. 10, pp. 3019–3041.

Jun, S., Park, D. and Yeom, J. (2014), "The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference", *Technological Forecasting and Social Change*, Elsevier Inc., Vol. 86, pp. 237–253.

Krumeich, J., Werth, D. and Loos, P. (2016), "Prescriptive Control of Business Processes", *Business & Information Systems Engineering*, Springer Fachmedien Wiesbaden, Vol. 58 No. 4, pp. 261–280.

Kumar, A., Shankar, R., Choudhary, A. and Thakur, L.S. (2016), “A big data MapReduce framework for fault diagnosis in cloud-based manufacturing”, *International Journal of Production Research*, Vol. 54 No. 23, pp. 7060–7073.

Kuo, R.J., Pai, C.M., Lin, R.H. and Chu, H.C. (2015), “The integration of association rule mining and artificial immune network for supplier selection and order quantity allocation”, *Applied Mathematics and Computation*, Elsevier Inc., Vol. 250, pp. 958–972.

Lee, C.K.H. (2016), “A GA-based optimisation model for big data analytics supporting anticipatory shipping in Retail 4.0”, *International Journal of Production Research*, Vol. 54 No. August, pp. 1–13.

Lee, C.K.H., Choy, K.L., Ho, G.T.S. and Lin, C. (2016), “A cloud-based responsive replenishment system in a franchise business model using a fuzzy logic approach”, *Expert Systems*, Vol. 33 No. 1, pp. 14–29.

Lei, N. and Moon, S.K. (2015), “A Decision Support System for market-driven product positioning and design”, *Decision Support Systems*, Elsevier B.V., Vol. 69, pp. 82–91.

Li, B., Ch’ng, E., Chong, A.Y. and Bao, H. (2016), “Predicting online e-marketplace sales performances: A big data approach”, *Computers & Industrial Engineering*, Elsevier Ltd, Vol. Accepted, doi:10.1016/j.cie.2016.08.009.

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D. and Hampapur, A. (2014), “Improving rail network velocity: A machine learning approach to predictive maintenance”, *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 45, pp. 17–26.

Li, J., Moghaddam, M. and Nof, S.Y. (2016), “Dynamic storage assignment with product affinity and ABC classification—a case study”, *The International Journal of Advanced Manufacturing Technology*, The International Journal of Advanced Manufacturing

Technology, Vol. 84 No. 9–12, pp. 2179–2194.

Li, L., Su, X., Wang, Y., Lin, Y., Li, Z. and Li, Y. (2015), "Robust causal dependence mining in big data network and its application to traffic flow predictions", *Transportation Research Part C: Emerging Technologies*, Vol. 58, pp. 292–307.

Li, X., Song, J. and Huang, B. (2016), "A scientific workflow management system architecture and its scheduling based on cloud service platform for manufacturing big data analytics", *The International Journal of Advanced Manufacturing Technology*, Vol. 84 No. 1–4, pp. 119–131.

Ling Ho, C. and Wen Shih, H. (2014), "Applying Data Ming to Develop a Warning System of Procurement in Construction", *International Journal of Future Computer and Communication*, Vol. 3 No. 3, pp. 168–171.

Ma, J., Kwak, M. and Kim, H.M. (2014), "Demand Trend Mining for Predictive Life Cycle Design", *Journal of Cleaner Production*, Elsevier Ltd, Vol. 68, pp. 189–199.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011), *Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute*, doi:10.1080/01443610903114527.

Marine-Roig, E. and Anton Clavé, S. (2015), "Tourism analytics with massive user-generated content: A case study of Barcelona", *Journal of Destination Marketing & Management*, Elsevier, Vol. 4 No. 3, pp. 162–172.

Mayring, P. (2008), *Qualitative Inhaltsanalyse (Qualitative content analysis)*, Weinheim, Beltz, 10thed.

Mehmood, R., Meriton, R., Graham, G., Hennelly, P., Kumar, M., Mehmood, R., Meriton, R., et al. (2017), "Exploring the influence of big data on city transport operations : a Markovian approach", *Journal of Operations & Production Management*, Vol. 37 No. 1, pp. 75–104.

Meziane, F. and Proudlove, N. (2000), “Intelligent systems in manufacturing : current developments and future prospects”, *Integrated Manufacturing Systems,* Vol. 11 No. 4, pp. 218–238.

Miroslav, M., Miloš, M., Velimir, Š., Božo, D. and Đorđe, L. (2014), “Semantic technologies on the mission: Preventing corruption in public procurement”, *Computers in Industry,* Vol. 65 No. 5, pp. 878–890.

Mishra, D., Gunasekaran, A., Papadopoulos, T., & Childe, S. J. (2016), "Big Data and supply chain management: a review and bibliometric analysis", *Annals of Operations Research*, 1–24.

Mori, J., Kajikawa, Y., Kashima, H. and Sakata, I. (2012), “Machine learning approach for finding business partners and building reciprocal relationships”, *Expert Systems with Applications,* Elsevier Ltd, Vol. 39 No. 12, pp. 10402–10407.

O’Donovan, P., Leahy, K., Bruton, K. and O’Sullivan, D.T.J. (2015), “Big data in manufacturing: a systematic mapping study”, *Journal of Big Data*, Journal of Big Data, Vol. 2 No. 1, p. 20.

Olson, D.L. (2015), “A Review of Supply Chain Data Mining Publications”, *Journal of Supply Chain Management Science*, Vol. 9, pp. 1–13.

Ong, J.B.S., Wang, Z., Goh, R.S.M., Yin, X.F., Xin, X. and Fu, X. (2015), “Understanding Natural Disasters as Risks in Supply Chain Management through Web Data Analysis”, *International Journal of Computer and Communication Engineering*, Vol. 4 No. 2, pp. 126–133.

Opresnik, D. and Taisch, M. (2015), “The value of Big Data in servitization”, *International Journal of Production Economics*, Elsevier, Vol. 165, pp. 174–184.

Oracle. (2012), *Big Data for the Enterprise*.

Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S.J. and Fosso-Wamba, S.

(2017), “The role of Big Data in explaining disaster resilience in supply chains for sustainability”, *Journal of Cleaner Production*, Elsevier Ltd, Vol. 142, pp. 1108–1118.

Prasad, S., Zakaria, R. and Altay, N. (2016), “Big data in humanitarian supply chain networks: a resource dependence perspective”, *Annals of Operations Research*, Springer US, Vol. August, pp. 1–31.

Rehman, M.H. ur, Chang, V., Batool, A. and Wah, T.Y. (2016), “Big data reduction framework for value creation in sustainable enterprises”, *International Journal of Information Management*, Elsevier Ltd, Vol. 36 No. 6, pp. 917–928.

Ramanathan, U., Subramanian, N., Parrott, G. (2017) "Role of social media in retail network operations and marketing to enhance customer satisfaction", *International Journal of Operations & Production Management*, Vol. 37 No. 1, pp.105–123.

Rozados, I.V. and Tjahjono, B. (2014), “Big Data Analytics in Supply Chain Management: Trends and Related Research”, *6th International Conference on Operations and Supply Chain Management*, Bali.

Salehan, M. and Kim, D.J. (2016), “Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics”, *Decision Support Systems*, Elsevier B.V., Vol. 81, pp. 30–40.

Sanders, N.R. (2014), *Big Data Driven Supply Chain Management*, Pearson Education, Inc., available at: http://ptgmedia.pearsoncmg.com/images/9780133801286/samplepages/0133801284.pdf

Saumyadipta, B.L.S.P., Rao, P. and Rao, S.B. (2016), *Big Data Analytics: Methods and Applications*, Springer, available at: 5. https://books.google.co.uk/books?id=_xhADQAAQBAJ&printsec=frontcover#v=onepage&q=prescriptive&f=false.

Schmidt, B., Flannery, P. and DeSantis, M. (2015), “Real-Time Predictive Analytics, Big

Data & Energy Market Efficiency: Key to Efficient Markets and Lower Prices for Consumers", *Applied Mechanics and Materials*, Vol. 704, pp. 453–458.

Schoenherr, T., & Speier- Pero, C. (2015), "Data science, predictive analytics, and big data in supply chain management: Current state and future potential", *Journal of Business Logistics*, Vol. 36, No. 1, pp. 120–132.

Seuring, S. (2013), "A review of modeling approaches for sustainable supply chain management", *Decision Support Systems*, Elsevier B.V., Vol. 54 No. 4, pp. 1513–1520.

Seuring, S. and Müller, M. (2008), "From a literature review to a conceptual framework for sustainable supply chain management", *Journal of Cleaner Production*, Vol. 16 No. 15, pp. 1699–1710.

Shan, Z. and Zhu, Q. (2015), "Camera location for real-time traffic state estimation in urban road network using big GPS data", *Neurocomputing*, Elsevier, Vol. 169, pp. 134–143.

Sheffi, Y. (2015), "Preparing for Disruptions Through Early Detection Preparing for Disruptions Through Early Detection", *MIT Sloan Management Review*, Vol. 57 No. 1, pp. 31–42.

Shi, Q. and Abdel-Aty, M. (2015), "Big Data applications in real-time traffic operation and safety monitoring and improvement on urban expressways", *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 58, pp. 380–394.

Shu, Y., Ming, L., Cheng, F., Zhang, Z. and Zhao, J. (2016), "Abnormal situation management: Challenges and opportunities in the big data era", *Computers & Chemical Engineering*, Elsevier Ltd, Vol. 91, pp. 104–113.

Sivamani, S., Kwak, K. and Cho, Y. (2014), "A Study on Intelligent User-Centric Logistics Service Model Using Ontology", *Journal of Applied Mathematics*, Vol. 2014 No. 162838, pp. 1–10.

Song, M.-L., Fisher, R., Wang, J.-L. and Cui, L.-B. (2016), "Environmental performance evaluation with big data: theories and methods”, *Annals of Operations Research*, Springer US, Vol. March, pp. 1–14.

St-Aubin, P., Saunier, N. and Miranda-Moreno, L. (2015), “Large-scale automated proactive road safety analysis using video data”, *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 58, pp. 363–379.

Stefanovic, N. (2015), “Collaborative predictive business intelligence model for spare parts inventory replenishment”, *Computer Science and Information Systems*, Vol. 12 No. 3, pp. 911–930.

Tan, K.H., Zhan, Y., Ji, G., Ye, F. and Chang, C. (2015), “Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph”, *International Journal of Production Economics*, Elsevier, Vol. 165, pp. 223–233.

Tan, M. and Lee, W. (2015), “Evaluation and Improvement of Procurement Process with Data Analytics”, *International Journal of Advanced Computer Science and Applications*, Vol. 6 No. 8, pp. 70–80.

Ting, S.L.L., Tse, Y.K.K., Ho, G.T.S.T.S., Chung, S.H.H. and Pang, G. (2014), “Mining logistics data to assure the quality in a sustainable food supply chain: A case in the red wine industry”, *International Journal of Production Economics*, Vol. 152, pp. 200–209.

Toole, J.L., Colak, S., Sturt, B., Alexander, L.P., Evsukoff, A. and González, M.C. (2015), “The path most traveled: Travel demand estimation using big data resources”, *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 58, pp. 162–177.

Tranfield, D., Denyer, D. and Smart, P. (2003), “Towards a methodology for developing evidence-informed management knowledge by means of systematic review””, *British Journal of Management*, Vol. 14 No. 3, pp. 207–222.

Tsai, C.-W., Lai, C.-F., Chao, H.-C. and Vasilakos, A. V. (2015), “Big data analytics: a survey”, *Journal of Big Data*, Springer International Publishing, Vol. 2 No. 1, p. 21.

Tsai, C.-Y. and Huang, S.-H. (2015), “A data mining approach to optimise shelf space allocation in consideration of customer purchase and moving behaviours”, *International Journal of Production Research*, Vol. 53 No. 3, pp. 850–866.

Tu, W., Li, Q., Fang, Z., Shaw, S., Zhou, B. and Chang, X. (2016), “Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach”, *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 65 No. 3688, pp. 172–189.

Waller, M. A., and Fawcett, S. E. (2013). "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management". *Journal of Business Logistics*, 34(2), 77–84.

Walker, G. and Strathie, A. (2016), “Big data and ergonomics methods: A new paradigm for tackling strategic transport safety risks”, *Applied Ergonomics*, Elsevier Ltd, Vol. 53, pp. 298–311.

Wamba, S.F., Akter, S., Coltman, T. and Ngai, E.W.T. (2015), “Information technology-enabled supply chain management”, *Production Planning & Control*, Vol. 26 No. 12, pp. 933–944.

Wamba, S.F., Akter, S., Edwards, A., Chopin, G. and 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*, Elsevier, Vol. 165 No. July, pp. 234–246.

Wamba, S. F., Ngai, E.W.T., Riggins, F. and Akter, S. (2017), "Transforming operations and production management using big data and business analytics: future research directions", *International Journal of Operations & Production Management*, Vol. 37

No. 1, pp. 2–9.

Wang, C., Li, X., Zhou, X., Wang, A. and Nedjah, N. (2016), “Soft computing in big data intelligent transportation systems”, *Applied Soft Computing*, Vol. 38, pp. 1099–1108.

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

Wang, J. and Zhang, J. (2016), “Big data analytics for forecasting cycle time in semiconductor wafer fabrication system”, *International Journal of Production Research*, Vol. 54 No. 23, pp. 7231–7244.

Wang, J., Zhang, L., Duan, L. and Gao, R.X. (2015), “A new paradigm of cloud-based predictive maintenance for intelligent manufacturing”, *Journal of Intelligent Manufacturing*, Vol. March, pp. 1–13.

Wang, S., Wan, J., Zhang, D., Li, D. and Zhang, C. (2016), “Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination”, *Computer Networks*, Vol. 101, pp. 158–168.

Wang, Z., Tu, L., Guo, Z., Yang, L.T. and Huang, B. (2014), “Analysis of user behaviors by mining large network data sets”, *Future Generation Computer Systems*, Elsevier B.V., Vol. 37, pp. 429–437.

Wegner, M. and Küchelhaus, M. (2013), *Big Data in Logistics*, *DHL Customer Solutions & Innovation*.

White, M. (2012), “Digital workplaces: vision and reality”, *Business Information Review*, Vol. 29 No. 4, pp. 205–214.

Wu, K., Liao, C., Tseng, M., Lim, M.K., Hu, J. and Tan, K. (2017), “Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties”, *Journal of Cleaner Production*, Elsevier Ltd, Vol. 142, pp. 663–676.

Xia, D., Wang, B., Li, H., Li, Y. and Zhang, Z. (2016), "A distributed spatial–temporal weighted model on MapReduce for short-term traffic flow forecasting", *Neurocomputing*, Elsevier, Vol. 179 No. ii, pp. 246–263.

Yan-Qiu, L. and Hao, W. (2016), "Optimization for service supply network base on the user's delivery time under the background of big data", *2016 Chinese Control and Decision Conference (CCDC)*, IEEE, Vol. 13, pp. 4564–4569.

Yu, R. and Abdel-Aty, M. (2014), "Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data", *Safety Science*, Elsevier Ltd, Vol. 63, pp. 50–56.

Zangenehpour, S., Miranda-Moreno, L.F. and Saunier, N. (2015), "Automated classification based on video data at intersections with heavy pedestrian and bicycle traffic: Methodology and application", *Transportation Research Part C: Emerging Technologies*, Elsevier Ltd, Vol. 56, pp. 161–176.

Zhang, C., Yao, X. and Zhang, J. (2015), "Abnormal Condition Monitoring of Workpieces Based on RFID for Wisdom Manufacturing Workshops", *Sensors*, Vol. 15 No. 12, pp. 30165–30186.

Zhang, J., Meng, W., Liu, Q., Jiang, H., Feng, Y. and Wang, G. (2016), "Efficient vehicles path planning algorithm based on taxi GPS big data", *Optik - International Journal for Light and Electron Optics*, Elsevier GmbH., Vol. 127 No. 5, pp. 2579–2585.

Zhang, Y., Ren, S., Liu, Y. and Si, S. (2017), "A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products", *Journal of Cleaner Production*, Elsevier Ltd, Vol. 142, pp. 626–641.

Zhang, Y., Zhang, G., Du, W., Wang, J., Ali, E. and Sun, S. (2015), "An optimization method for shopfloor material handling based on real-time and multi-source manufacturing data”, *International Journal of Production Economics*, Elsevier, Vol. 165, pp. 282–292.

Zhao, R., Liu, Y., Zhang, N. and Huang, T. (2016), “An optimization model for green supply chain management by using a big data analytic approach”, *Journal of Cleaner Production*, Elsevier Ltd, Vol. 142, pp. 1085–1097.

Zhao, X., Yeung, K., Huang, Q. and Song, X. (2015), “Improving the predictability of business failure of supply chain finance clients by using external big dataset”, (Xiaojun Wang, Professor Leroy White, D.,Ed.)*Industrial Management & Data Systems*, Vol. 115 No. 9, pp. 1683–1703.

Zhong, R.Y., Huang, G.Q., Lan, S., Dai, Q.Y., Chen, X. and Zhang, T. (2015), “A big data approach for logistics trajectory discovery from RFID-enabled production data”, *International Journal of Production Economics*, Vol. 165, pp. 260–272.

Zhong, R.Y., Huang, G.Q., Lan, S., Dai, Q.Y., Zhang, T. and Xu, C. (2015), “A two-level advanced production planning and scheduling model for RFID-enabled ubiquitous manufacturing”, *Advanced Engineering Informatics*, Elsevier Ltd, Vol. 29 No. 4, pp. 799–812.

Zhong, R.Y., Lan, S., Xu, C., Dai, Q. and Huang, G.Q. (2016), “Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing”, *The International Journal of Advanced Manufacturing Technology*, Vol. 84 No. 1–4, pp. 5–16.

Zhong, R.Y., Xu, C., Chen, C. and Huang, G.Q. (2015), “Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors”, *International Journal of Production Research*, Vol. 7543 No. September, pp. 1–12.