Chapter 7
Big Data Analytics and Its Applications in Supply Chain Management: A Literature Review Using SCOR Model

Abhijeet Ghadge and D. G. Mogale

Abstract As we move further into the twenty-first century, technology has played an essential role in people's lives inevitably, and generated massive information called 'Big Data (BD)'. With the ability to manage massive dataset, big data analytics can usefully extract the insight from big data and support the firm to leverage decision-making. Hence, the interest in big data applications has spread to comprehend many areas of study, including Supply Chain Management (SCM). However, the academic articles studying the employment of Big Data Analytics (BDA) in SCM are limited. Besides, most of those academic papers offer less interest in the entire SC systems. Most of them prefer to study in an individual SC area. Thus, this study aims to investigate state of the art in this domain through Systematic Literature Review (SLR) and discuss future research opportunities. We found that optimisation, simulation and visualisation tend to be the most appropriate BD tools to apply in SCM. Also, linear programming, statistics, association rule mining, fuzzy logic and decision tree are likely to be the most suitable BDA techniques for SC operations.

Keywords Big data analytics · Supply chain management · Big data tools and techniques · SCOR model · And Systematic literature review

7.1 Introduction

Big data (BD) is mostly interpreted as the enormous collection of data that has rapidly increased ahead of the analytical capability of traditional applications. Consequently, Dey et al. [6] described that big data is the massive store of information kept on any

A. Ghadge
Centre for Logistics and Supply Chain Management, School of Management, Cranfield University, Cranfield MK43 0AL, UK
e-mail: Abhijeet.Ghadge@cranfield.ac.uk

D. G. Mogale (✉)
Logistics and Operations Management Section, Cardiff Business School, Cardiff University, Cardiff, UK
e-mail: MogaleD@cardiff.ac.uk

M. K. Tiwari et al. (eds.), *Applications of Emerging Technologies and AI/ML Algorithms*, Asset Analytics, https://doi.org/10.1007/978-981-99-1019-9_7

servers that is difficult to analyse by conventional tools. Even definitions of BD generally concentrate on data size, other characteristics comprehensively explain the word 'big data' known as the four Vs: Volume—the size of data, Velocity—the speed of data, Variety—the structure form of data and Veracity—the accuracy of data. BD presents characteristically as large volume, high velocity, less structured or unstructured forms, and less accuracy.

Big data analytics (BDA) has a high capability to manage the complexity of BD. Bolstorff and Rosenbaum [3] explained that BDA includes 'big data' and 'advanced analytics' which can discover valuable business trends. Arunachalam et al. [1] classified data analytics techniques into Descriptive, Predictive and Prescriptive analytics. Supply Chain Operations Reference (SCOR), established by the Supply Chain Council (SCC), is a reference model for SC activities to facilitate the SC organisations in case of exchanging the knowledge or information. The SCOR model includes five compositions (Plan, Source, Make, Deliver and Return) [7]. Presently, limited authors discussed the collaboration of BDA and SCM and they explained specific operation like production, planning and procurement [12]. Hence, there is a need of holistic analysis of application of BDA in SC. Following this background, this study attempts to answer the research question: which are BDA tools and techniques that appropriate for each stage of the SCOR model? The aim of this study is to analyse the tools and techniques of BDA used in SCM and their benefits through a Systematic Literature Review (SLR).

The reminder of the paper is arranged as follows. Section 7.2 'Research Methodology' provides the methods of research delimitation. Then, the paper describes the key findings separated into 'descriptive findings' performed in Sect. 7.3 and 'thematic findings' described in Sect. 7.4. Lastly, Sect. 7.5 'conclusion' represents the research limitations and concluding statements.

7.2 Research Methodology

A literature review offers an essential tool for consolidating, improving, and synthesising several papers to create a research theory of study field [16]. According to Bearman et al. [2], the literature review can be classified into a traditional narrative review and a systematic review. The first one, the traditional review performs a specific literature review through the viewpoint of the reviewer. Secondly, the systematic literature review (SLR) is a descriptive approach that employs reproducible scientific procedures, such as the developed search strategies and a well-organised presentation of the research findings. Owing to Tranfield et al. [19], the procedure of SLR has been claimed to be the most efficient method that contributes to the high-quality results of the literature review. Based on Ghadge et al. [8], the following are three significant steps of the SLR approach: 1. Identification of sources, 2. Data screening and synthesis and 3. Data analysis and dissemination.

In the first process, search strings were listed through the exploration of the primary studies about the application of BDA in SCM. This procedure was operated based on the 'Boolean method'. Accordingly, the words, 'Big Data Analytics' and 'Supply Chain Management' were the main keywords. Also, the other words that frequently occur in the BDA and SCM literature were included in the searching words to cover all relevant fields of BDA and SCM. Then, two renowned academic databases, Scopus and Web of Science, were chosen for data extraction. Both data sources are commonly used for literature review and proven to offer inclusive outcomes [10].

There are five criteria applied for this study. The first one is the publication year. Secondly, language is identified to be one of the criteria. Only the literature in English were considered in this paper because English is a universal language for academic articles. Thirdly, only academic journals were selected for the study to ensure the quality of the included literature. Other academic and non-academic documents like letters, textbooks, book chapters, conference papers and editorial were excluded from the study. Fourthly, the comprehended articles of this research were peer-reviewed because these peer-reviewed literature have been guaranteed the journal quality by the experts. Finally, the full-text assessment may be significant for the literature review. It can confirm that the author can read all contents in the paper and synthesise the research method and findings comfortably. Then, all included papers of the study were full-text articles. In second stage, the inclusion and exclusion criteria were employed to screen the articles to obtain the high quality of the evidence-based approach. Only literature that met all the pre-defined criteria were included in the literature review [19]. Accordingly, a framework for data screening and synthesis was created based on inclusion and exclusion criteria. After the data extraction and the criteria examination, the authors read the title, keywords and abstracts of each paper to approve that their contents involve the application of BDA in SCM. The studies that cover only BDA or SCM individually were excluded from the lists of selected journals.

In the last stage, the method of data analysis and dissemination is recognised as the most meaningful section of SLR because this part presents the summary and evaluation of article findings that developed from the collective literature to the readers [19]. Overall, data analysis is split into the descriptive and thematic analysis. The descriptive analysis applies statistical knowledge to explain the results. For the thematic analysis, the information from the included literature review will be developed to achieve the research objective and answer the research question.

7.3 Descriptive Analysis

A descriptive finding of previous studies obtained from Scopus and Web of science following the SLR process in Sect. 7.2 is presented here. The results of a systematic review are statistically described owing to publication year. There are 20

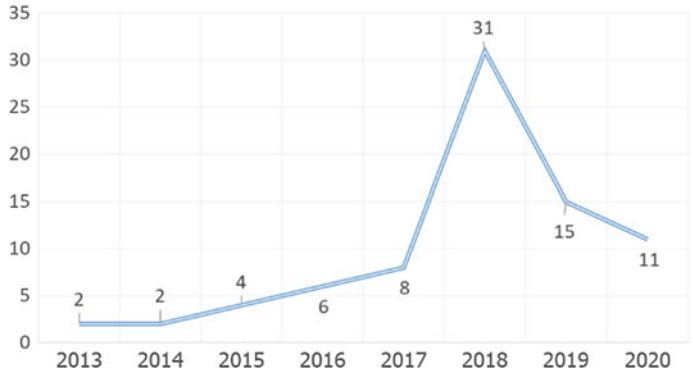

Fig. 7.1 The yearly number of BDA articles in the context of SCM

published journals distributed from the included works of literature. Further, the articles published in the International Journal of Production Research and Computer & Industrial Engineering mostly appear in the literature review, eight papers of each journal.

7.3.1 Year Wise Publications

Figure 7.1 shows the yearly number of studies increases continuously, and peak in 2018, 31 articles. Then, the quantities of papers decrease to 15 in 2019, and 11 in the half-year of 2020.

7.3.2 Methodological Distribution of Articles

Based on Fig. 7.2, the quantitative method, 49%, is the most favoured approach used by the researchers to conduct the academic studies related to BDA and SCM, followed by the qualitative method, 42%. Also, a mixed method which employs both qualitative and quantitative methodology offers the least popular practice among the authors, 9%.

7.3.3 Distribution of BDA Types

According to Fig. 7.3, predictive analytics that supports future forecasting is the most prevalent BDA type when applying BDA in SCM, 43%. This statistical result agrees with the descriptive conclusion of supply chain function and SCOR activities that

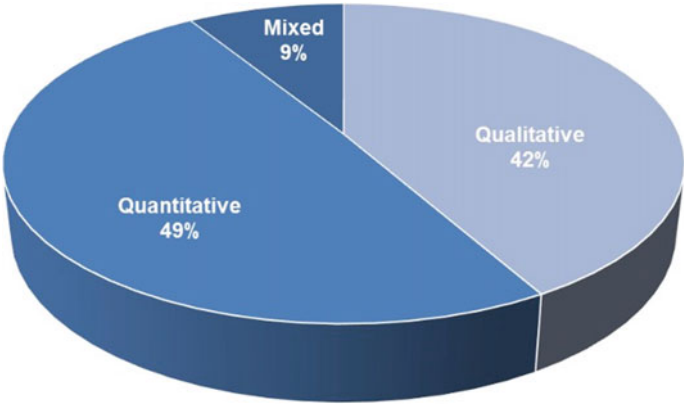

Fig. 7.2 The percentage of articles based on the methodologies

retail/customer and plan process are the preferable areas to conduct BDA. Further, prescriptive places as the second rank of BDA types that usually applies in SCM research with 31%, while descriptive/diagnostic shows the least percentage, 26%.

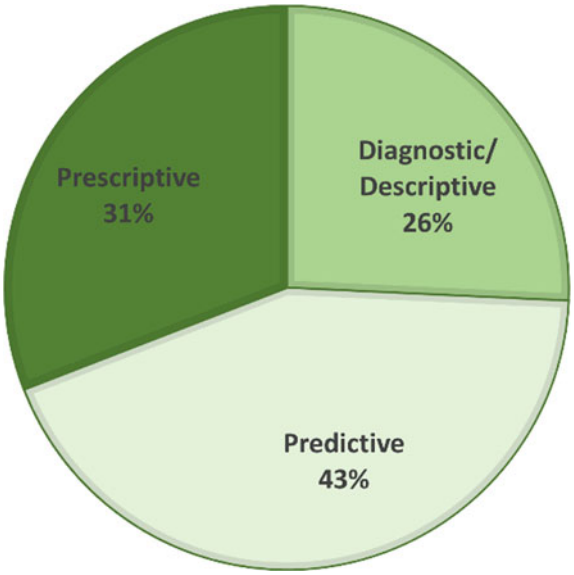

Fig. 7.3 The percentage of articles based on BDA types

7.3.4 Distribution of BDA Tools

Owing to Fig. 7.4, mixed technique, especially data mining, machine learning and statistical modelling, tend to be the most popular tools, 19.67%, used to solve SCM issues. An example of an academic article that suggests the data mining approach is 'A data mining-based framework for supply chain risk management [13]'. Also, the instance research that recommends the machine learning technique is 'Supply chain management and Industry 4.0: conducting research in the digital age [11]'.

7.4 Thematic Analysis

Thematic analysis aims to answer the research question.

SCOR model activities: Theme to describe the elements of the SCOR model, which consists of plan, source, make, deliver, and return.

Supply chain functions: Theme to understand the core events of logistics and supply chain management comprising procurement, manufacturing, logistics and transportation, warehousing and distribution, and retail and customer. The theme is established based on Nguyen et al. [15] to explore which SCM functions relate to each SCOR activity.

Types of BDA analytics: Theme to provide the classification of big data models is categorised into descriptive and diagnostics analytics, predictive analytics and prescriptive analytics. This theme is developed by referencing Arunachalam et al. [1] to explain the level of data analytics applied in a conceptual model.

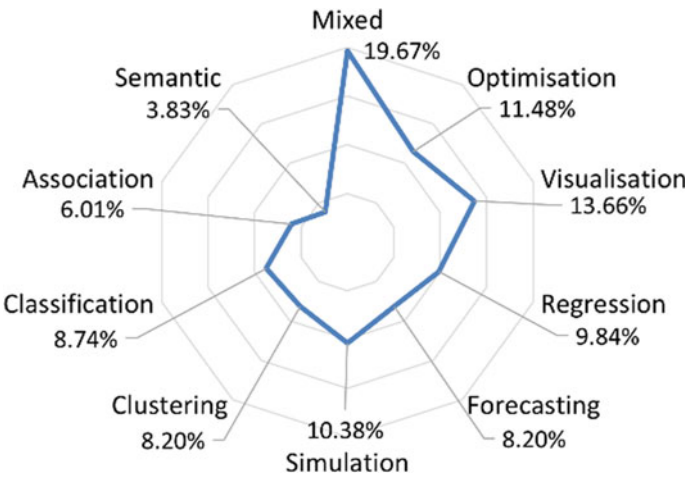

Fig. 7.4 The comparison of BDA tools applied in SCM

BDA tools: Theme to indicate the BDA tools is employed in SCM. The subcategories of this theme are identified according to Nguyen et al. [15] which include association, clustering, classification, regression, forecasting, semantic, optimisation, simulation, visualisation and mixed tools. For mixed algorithms, the author discovers from the descriptive analysis that data mining, machine learning and basic statistical modelling mostly present in the selected articles. Then, these three algorithms are added to the BDA tools to increase the accuracy of data analysis. The appropriate BDA tools are chosen based on the frequency of appearance in the literature review.

BDA techniques: After the literature review, the theme of BDA techniques is created to explore the most suitable BDA techniques used in SCM. The subcategories of this theme are essentially identified according to the findings from SLR, Nguyen et al. [15] and Chehbi-Gamoura et al. [5] including association rule mining (ARM), spatial-temporal visualisation (STV), decision tree (DT), support vector machine (SVM), linear programming (LP), linear regression (LIR), logistics regression (LOG), genetic algorithm (GA), neural network (NN), text mining (TM), k-mean clustering (KMC), time-series forecasting (TSF), Bayesian network (BN), fuzzy logic (FL), sequential analysis (SEG), sentiment analysis (SEN), natural language processing (NLP) and statistics (STA). The proposed BDA techniques are selected due to the frequency of presence in the SLR.

7.4.1 Plan Process

According to Wang et al. [20], BDA provides a significant role in the planning stage of SCM. It encourages organisations to enhance their decision making on demand planning, sourcing, production planning, inventory management, transportation, and the supply chain network design. Then, it is noticeable that the use of BDA in the planning phase is likely to cover all critical events in the supply chain.

Tiwari et al. [18] state that big data applications can notice demand signal, which facilitates the companies in terms of demand forecasting. Hence, organisations are likely to caption the new trends of markets and prepare for the change in the market and the fluctuation of demand more effectively. Consequently, BDA can support planners through the tracking of cash-cycle management, inventory replenishment and forecasting accuracy by tracking day-to-day sales information. Thereby, organisations can prevent missing products which lead to lost sales and good preparation for the peak season of products Hahn and Packowski [9]. Thus, BDA tends to offer improvements to supply chain planning significantly. In order to provide more specific guidance on the BDA employment in SCM, this study has explored the most appropriate BDA tools and techniques for the planning stage through the literature review.

7.4.2 Source Process

According to Nguyen et al. [15] and Tiwari et al. [18], the use of BDA in procurement frequently comprehends supplier selection, sourcing improvement, supply risk management and supplier performance. With the aim of leveraging supply risk management and supplier performance, big data applications can differentiate risks into the risk that should avoid and risk that should take to monitor market trends. Thereby, the firms can update and capture the alterations in sourcing or supplier markets, helping the company to respond to the supply change capably [20]. For supplier selection and sourcing improvement, BDA can efficiently investigate and analyse supplier performance [20]. Further, BDA can offer in-depth data analysis of the return on investment (ROI). Hence, the organisations may improve strategic sourcing through the worthy of financial information Tiwari et al. [18]. In an attempt to observe more BDA utilisation in the sourcing supply chain, this research has examined the most suitable BDA tools and techniques that mostly apply in sourcing activity.

7.4.3 Make Process

Due to Nguyen et al. [15] production control currently gains the interest of applying BDA. Wang et al. [20] state that big data tools can describe the production costs to the manufacturers. Moreover, BDA can provide the impact of manufacturing profits when there are any changes in production costs. Accordingly, the big data techniques can ensure the right production parts go to the right assembly by extracting the insight of production capacity. Further, BDA can be used to examine the amount of material waste and seek the proper manufacturing tools to decrease the level of material waste.

As well as that, inventory management seems to be a significant area for applying BDA. Big data applications assist the companies in leveraging the design of current stock operations by using the optimisation system. It is likely to help the firms to decrease the challenges of multi-channel inventories [20]. Also, Tiwari et al. [18] suggest that the sharing of information on internal and external production system would enhance the benefits of big data on inventory management. Accordingly, this study has examined the most advisable big data tools and techniques that mostly apply in the production process.

7.4.4 Deliver Process

The logistics function relates to the distribution of goods from the supply sites to the distribution centres, and the end customers. Also, logistics data can be massively generated by many players in the delivery route, such as carriers, transportation

service providers, and shippers. Thereby, the development of BDA in distributing area seems significant to leverage SCM. The employment of BDA in logistics and transportation focuses on the fundamental logistics functions which consist of supply chain network design, route optimisation, and proactive safety management [17, 21]. With the purpose of enhancing the efficiency of supply chain network design, big data tools can calculate the optimal locations of the distribution centres by analysing the massive data generated from warehousing, transportation, and consumer demand [21].

7.4.5 Return Process

According to Chehbi-Gamoura et al. [5], the return process significantly applies big data applications in the areas of the green supply chain which specially concerns with ecology and cost improvement, reverse logistics management and the closed-loop supply chain. The green supply chain essentially focusses on the decrease in waste. Then, the returns of goods for remanufacturing, recycling and reuse can reduce the number of wastes spread in the environment. To support the green supply chain, BDA can be applied in the process control to restrict the pollution emission and manage the natural resources sustainably [21].

Furthermore, big data applications have a high potential to employ in reverse logistics and closed-loop supply chains. Nguyen et al. [15] claim that BDA has been used in product lifecycle design which is useful for enterprises to forecast the number of returned products. This information is significantly beneficial for the planning of production capacity and remanufacturing scheduling in reverse logistics. Consequently, big data enables organisations to adopt customer complaints to identify the root causes of defects. It enables the company to improve its product quality [4]. Optimisation and visualisation are likely to be the most appropriate BDA tools for raising the performance of the return process.

7.5 Conclusion and Future Work

Following the SLR process, this study aims to examine the state of the art in the domain of big data applications in SCM based on the SCOR model. Consequently, the study demonstrates BDA tools and techniques that seem appropriate to apply in each stage of the SCOR model. With an attempt to achieve the goal, this study aims to answer the research question: which are big data analytical tools and techniques that appropriate for each of the SCOR models?

In order to address research question, the SLR framework is conducted to obtain a high quality of academic articles for the literature review. Accordingly, papers related to the application of BDA in the context of supply chain operations are selected to collect and analyse critical information. The descriptive analysis shows

that the researchers have significantly gained an interest in the utilisation of BDA in SCM during the last four years. Consequently, most of the researchers seem to concentrate on the strategic level of SCM. As well as that, demand management is related to retail, and customer tends to be the most favoured area for applying big data technology.

According to the thematic analysis, it has discovered the appropriate BDA tools and techniques for each stage of the SCOR model, which answer the research question potentially. The results show that LP optimisation, and ARM, DT, and FL simulations have a high potential to improve planning performance. Also, STA visualisation, LP optimisation, and ARM and DT simulations tend to support the operations in procurement usefully. To better the manufacturing practice, STA visualisation, LP and STA optimisations, and FL and STA simulations can be employed in this field. Consequently, LP optimisation, STA visualisation and STA simulation can be applied in logistics and transportation management to develop the delivery functions. For return activity, STA optimisation and STA visualisation are likely to leverage the reverse logistics performance.

The study offers significant implications for the theoretical research overview. It introduces the potential BDA algorithms that have a high capacity to improve the entire SCM with further academic evidence. The findings also provide the most suitable BDA tools and techniques to apply in the individual stage of SCM owing to the SCOR model and level of management. These findings tend to support the statement of Chehbi-Gamoura et al. [5] that the study of the employment of BDA across the whole supply chain at one research is essentially needed to manage the massive information generated in the SCM simultaneously.

The results of this study provide several practical implications for the practitioners. First, this study offers the guidance of what BDA tools and techniques that the organisations should consider applying in their supply chain practice to leverage the performance. The integrated big data applications from the beginning to the ending stages of SCM are presented through the conceptual model. It would support the company on the decision of big data investment across the entire chain.

Even this study is likely to offer useful insights, there are some limitations to concern. Firstly, the author only extracted the academic literature from SCOPUS and WOS databases to explore the research implications. Hence, some relevant articles published on other sources can be missed from the literature review. Also, the results are determined based on the author perspective only. Therefore, some analysis may be presented in different ways compared to the other papers. Thus, future research should employ the cross-checking process to minimise bias and errors []. A conceptual model integrating different BDA tools and techniques could be developed for guidance for the researchers or practitioners in case of BDA investment in SCM.

References

1. Arunachalam D, Kumar N, Kawalek JP (2018) Understanding big data analytics capabilities in supply chain management: unravelling the issues, challenges and implications for practice. Transp Res Part E Logist Transp Rev 114:416–436
2. Bearman M, Smith CD, Carbone A, Slade S, Baik C, Hughes-Warrington M, Neumann DL (2012) Systematic review methodology in higher education. High Educ Res Dev 31(5):625–640
3. Bolstorff P, Rosenbaum R (2011) Supply chain excellence: a handbook for dramatic improvement using the SCOR model, 3rd edn.
4. Brinch M, Stentoft J, Jensen JK, Rajkumar C (2018) Practitioners understanding of big data and its applications in supply chain management. Int J Logist Manage.
5. Chehbi-Gamoura S, Derrouiche R, Damand D, Barth M (2020) Insights from big data analytics in supply chain management: an all-inclusive literature review using the SCOR model. Prod Plan Control 31(5):355–382
6. Dey N, Das H, Naik B, Behera HS (eds) (2019) Big data analytics for intelligent healthcare management. Academic Press
7. Georgise FB, Thoben KD, Seifert M (2012) Adapting the SCOR model to suit the different scenarios: a literature review & research agenda. Int J Bus Manage 7(6):2
8. Ghadge A, Wurtmann H, Seuring S (2020) Managing climate change risks in global supply chains: a review and research agenda. Int J Prod Res 58(1):44–64
9. Hahn GJ, Packowski J (2015) A perspective on applications of in-memory analytics in supply chain management. Decis Support Syst 76:45–52
10. Harzing AW, Alakangas S (2016) Google Scholar, Scopus and the Web of Science: a longitudinal and cross- disciplinary comparison. Scientometrics 106(2):787–804
11. Hofmann E, Sternberg H, Chen H, Pflaum A, Prockl G (2019) Supply chain management and Industry 4.0: conducting research in the digital age. Int J Phys Distrib Logistics Manage
12. Kache F, Seuring S (2017) Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. Int J Oper Prod Manage
13. Kara ME, Fırat SÜO, Ghadge A (2020) A data mining-based framework for supply chain risk management. Comput Ind Eng 139:105570
14. Mishra D, Gunasekaran A, Papadopoulos T, Childe SJ (2018) Big Data and supply chain management: a review and bibliometric analysis. Ann Oper Res 270(1):313–336
15. Nguyen T, Li ZHOU, Spiegler V, Ieromonachou P, Lin Y (2018) Big data analytics in supply chain management: a state-of-the-art literature review. Comput Oper Res 98:254–264
16. Seuring S, Gold S (2012) Conducting content-analysis based literature reviews in supply chain management. Supply Chain Manage Int J
17. Singh A, Kumari S, Malekpoor H, Mishra N (2018) Big data cloud computing framework for low carbon supplier selection in the beef supply chain. J Clean Prod 202:139–149
18. Tiwari S, Wee HM, Daryanto Y (2018) Big data analytics in supply chain management between 2010 and 2016: insights to industries. Comput Ind Eng 115:319–330
19. Tranfield D, Denyer D, Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br J Manag 14(3):207–222
20. Wang G, Gunasekaran A, Ngai EW, Papadopoulos T (2016) Big data analytics in logistics and supply chain management: certain investigations for research and applications. Int J Prod Econ 176:98–110
21. Zhao R, Liu Y, Zhang N, Huang T (2017) An optimization model for green supply chain management by using a big data analytic approach. J Clean Prod 142:1085–1097