Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2022, 9(10):75-82**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Supply Chain Analytics: Optimizing Operations with Big Data Insights

Pranay Mungara

Abstract Data science and big data analytics (DS &BDA) are vital to the supply chain and logistics sectors. But what little is known is spread out throughout a number of books and articles. A comprehensive and unbiased review strategy is necessary to bring order to the DS & BDA application areas in SC & L, which address paradigms of resilience, sustainability, and efficiency. Here, we offer a new angle on previous research on DS and BDA in SC and L, setting the stage for our analysis. Secondly, we categorize the present models/techniques used, organize their practical application areas, identify research gaps, and predict future research routes by using the methodologies presented for a comprehensive literature assessment on DS &BDA techniques in the SC &L domains. At different points in the decision-making process, recent review articles use a variety of DS and BDA-driven modelling approaches to analyse SC and L processes. Our research is based on a triangulation of three criteria: efficiency, resilience, and sustainability. This research examines the effects of big data analytics on critical supply chain processes. Though these concepts are widely used in supply chain management, there is little written about how the main processes work. As a result, this article offers a taxonomy of the existing literature on big data analytics' applications, as well as commentary from experts in the subject. In order to plan and execute a systematic literature mapping, we referred to a well-known collection of practical guidelines.

Keywords Data insights, supply chain analytics, optimization, big data

Introduction

Evolution in data generation and analysis are being brought about by advances in technology [1] asserts that big data have the potential to bring about changes in management as well as the entire method of conducting business. The concept of metamorphosis is thus extremely important for this reason. Big data is a resource for businesses that is valuable, diverse, fast, verifiable, and has five other properties. The present storage methods are being increasingly overwhelmed by the exponential growth of the amount of data, or "volume," which is measured in physical space needed to store it.

Both the intake and generation of data are impacted by the speed at which data are conveyed, as stated by the definition of speed provided by the same authors. According to [2], "variable data" is defined as information that can be derived from several sources. Since each platform has its unique data generation structure, there are no universally accepted standards.

Therefore, it is possible to generate data that is both partially ordered and completely organized. Furthermore, data relies on human judgment or is based on people's emotions, hence the term "veracity" is often used to describe how unreliable such data is. A definition like this was offered by IBM in 2012. Oracle supposedly established value in 2012 and associated it with the idea of essential data.

Comprehensive data analysis can yield high-value information, according to this. The veracity and practicality of the data make up the rigor of big data analytics (BDA), as stated in [3]. Data analysis is crucial to the other aspects of big data processing, which include volume, speed, and variety. Without it, these aspects would be meaningless. Improving decision-making capabilities is possible through the application of cutting-edge analytical methods that extract actionable insights from large data sets. Various technologies are being utilized in supply chain management (SCM) to link and coordinate all the links in the chain. These technologies include sensors, barcodes, RFIDs, and the internet of things. There have been a lot of studies and documents about how the BDA is changing supply networks due to its usage in supply chain management (SCM).

An increasing amount of research is pointing to the notion that using business process analysis (BDA) in supply chain management (SCM) can improve chain process agility, decrease operational costs, and identify the optimal areas to use BDA. To keep an eye on everything throughout the supply chain is what "supply chain management" is all about. According to Leveling [4], supply chains consist of companies whose primary activity is the transportation of commodities. The production and distribution of a product is an integrated process that involves many different companies working together in a supply chain.

Included in this chain are organizations that generate raw materials and components, as well as manufacturers of finished goods, distributors, retailers, and transportation providers. All participants in a supply chain agree that it is the network of businesses that coordinates the delivery of goods and services to customers. A number of different options are presented in Figure 1.

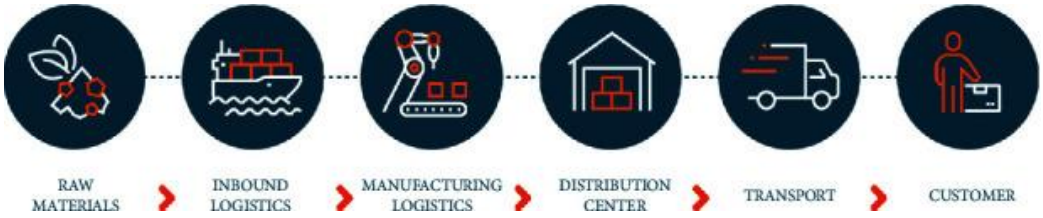

Figure 1: Supply chain

The author of [5] asserts that massive amounts of data are of no utility unless they can be interpreted. It is possible to assess the full value of big data when it is used to direct the decisions that businesses make. In order for organizations to be able to enable decision-making that is based on evidence, they need to be able to translate large amounts of data that is both diverse and fast into information that is valuable. There are five main steps to the process of using massive amounts of data to extract useful information. Following extraction, cleaning, and annotation, one should integrate, aggregate, and display the data. Then, one should model and analyze the data, and lastly, one should interpret the results.

The analysis of data management and the storage of information are the two primary subcategories that are separated from these five phases in [6]. For the purposes of analysis, the term "data management" refers to the procedures and instruments that are utilized to collect, store, prepare, and retrieve data. Alternatively, the term "analytics" relates to the methods that are utilized in order to investigate and get comprehension from significant data. The analysis of big data can be viewed as a minor stage inside the more complete process of extracting information from relevant data, as illustrated in Figure 2.

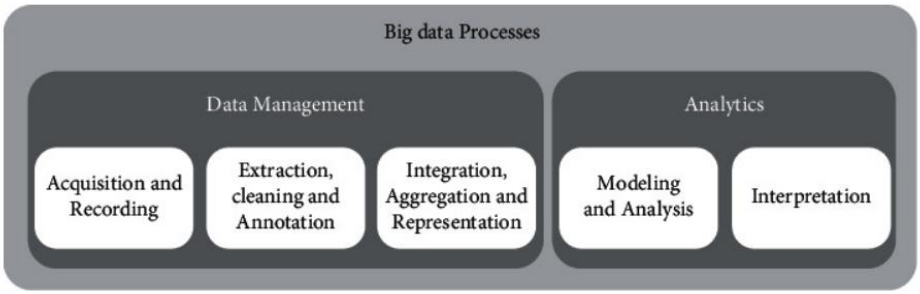

Figure 2 Big data processes

Literature Review

The various data sources accessible to supply chains (SCs) include enterprise resource planning (ERP) systems, sales, sales data, digital manufacturing, Blockchain, sensors, and client purchasing patterns [7,8].

There is an abundance of these data sets accessible. Depending on the circumstances, this material could be organized, partially organized, or completely unstructured. Utilizing big data analytics (BDA) to extract insights from data might enhance SC's performance and decision-making capacities.

The Business Data Analytics (BDA) presents organizations with major hurdles, despite the fact that it delivers substantial opportunity for value development [9]. Data science (DS) is concerned with more complicated data analytics, in contrast to business data analysis (BDA), which is concerned with the collection, storage, and analysis of data. In particular, methods for forecasting future outcomes are considered, such as deep learning algorithms and machine learning. According to the methodology, DS and BDA methods help with operational,

tactical, and strategic decisions in SC management. Businesses can get a commercial advantage by exploiting DS and BDA capabilities. Understanding client requests and needs, forecasting market changes, lowering costs, boosting sustainability, decreasing risk, and increasing resilience are just a few ways in which companies may enhance their SC design and management by utilizing DS and BDA methodologies [10].

In addition to the development of new methodologies, there has also been progress made in the area of DS and BDA tools. A better forecasting, optimization, and simulation model may be developed with the use of SC analytics software, which is beneficial to both researchers and practitioners. In addition, these tools are able to extract data and provide sophisticated visualizations. It is possible to build digital SC twins by integrating simulation and network design with SC operations data using SC-specific software like anyLogistixTM and LLamasoftTM [12]. These software programs are in addition to the enormous enterprises that are SAP®, IBM, and Oracle.

There is an increase in the number of chances available to SC researchers and practitioners as a result of modern methodological and software improvements. In contrast, there is a dearth of a systematic evaluation of the potential uses of DS and BDA in the SC &L domains, and the current insights are dispersed over several bodies of literature. We chose to do this thorough literature analysis in a systematic way since no previous reviews had adequately addressed the interplay between resilience, efficiency, and sustainability.

A. Data-driven optimisation

In recent years, the DDO has received a considerable degree of recognition. During the course of our investigation, we planned to find strategies that are interconnected with one another by incorporating the term "data-driven" in our list of keywords (see the preliminary search results for DDO in Table 1 for more information). In the field of mathematical programming, the DDO technique is a strategy that combines uncertainty-based optimization strategies with machine learning algorithms. There is a significant amount of time when the objective functions are connected to the costs [13]. In [14] employed four modeling approaches to separate DDO from other approaches. These methods of modeling were scenario-based optimization, resilient optimization, stochastic programming, and chance-constrained programming. Each of these modeling approaches was utilized to optimize the system.

It is possible that some of the problem criteria in the SC & L domain, such as consumer demand, manufacturing capacity, and delivery time, are regarded to be unclear. Directly inputs to the proposed programming issues are utilized by DDO approaches, which make use of knowledge regarding random variables. Traditional optimization models under uncertainty, on the other hand, take into account perfect knowledge for the parameters. This is in contrast to the situation described here.

Table 1: Initial search outcomes categorized by every set of keywords

Keywords	Database		
	Science direct	**Web of knowledge**	**Scopus**
Data science			
Supply chain	107	124	57
Logistics	57	34	111
Data-driven			
Supply chain	267	170	365
Logistics	89	87	876
Data mining			
Supply chain	138	152	873
Logistics	70	104	4157
Text mining			
Supply chain	20	35	86
Logistics	12	21	295

Keywords	Database		
	Science direct	Web of knowledge	Scopus
Big data			
Supply chain	203	666	1090
Logistics	78	287	1487
Data analytics			
Supply chain	323	267	415
Logistics	138	93	348
Predictive analytics			
Supply chain	70	198	109
Logistics	12	46	218
Machine learning			
Supply chain	73	133	414
Logistics	61	78	5937

Optimization under uncertainty was the subject of few surveys that made it into our literature review. The approaches that were most commonly used were stochastic programming techniques, such as MILP and MINLP. As an optimization technique, chance-constrained programming requires that all constraints be satisfied according to the probability distribution. The SC & L context is a good fit for this approach. For data sets to be considered for robust optimization, it is necessary to identify the uncertainty sets, which are sets of unknown parameters. To tackle the issue of ambiguity in the SC &L domain, this approach appears to be more effective. We mostly deal with ambiguous data, just like in the SC &L. Finding the best solution is the goal of scenario-based optimization, which makes use of uncertainty situations. Not a single one of the papers that we looked at used this strategy. Research into scenario-based data-driven optimization (DDO) approaches, particularly in risk management, appears to have promising prospects [15]. One possible future direction for DDO methodology in decision-making is the use of BDA techniques or technologies like cloud computing, parallel computing, Hadoop, Spark, or Map-Reduce solutions. The development of BDA applications to SC &L is ongoing, and this is especially true in this regard. Big data-driven optimization (BDDO) approaches utilize a combination of big data methods and strategies that leverage DDO. They have the potential to be fascinating for managing many SC & L issues.

Systematic Mapping Study

RQ1: The various components of the supply chain. For the purpose of finding a solution to this problem, it is necessary to determine which aspects of the supply chain make the most use of big data analytics. This region will give the key findings from Table 2. Of the fifty articles that were reviewed, 37 (or 78% of the total) dealt with demand and supply chain management. Since management and decision-making are continuously being engaged in these two domains, it is understandable that these results are concentrated there. In order to aid in decision-making in various domains, BDA techniques are utilized to raise awareness and provide data. Along with this data, knowing what the market wants has been a huge competitive advantage. Being good at this means keeping production under control and not making a ton of stuff when there isn't a need for it, which means less stock on hand and less money spent on materials and other frivolous things. In such instances, the inverse is also true: there is a great deal of market demand for relatively little production. Manufacturing, transportation/logistics, and storage/warehousing account for 22 percent, or 11 out of 50 articles analyzed. To optimize these areas, which are immediately related to production and aim to save time, cost, raw material, and other variables, BDA is often employed rather than focusing on management. This is because those areas are directly tied to production.

Table 2: SSMS research questions

Research questions	Motivation	Variable
RQ1: Where exactly is the application of big data analytics taking place within the supply chain management sector?	It is important to determine whether aspects of supply chain management are making substantial use of big data analytical methodologies.	Supply chain areas
RQ2: In these domains of supply chain management, what level of big data analytical analysis is utilized, and from what perspective?	Provide a list of the primary layers of big data analytics that are utilized for the management of supply chains.	Big data levels
RQ3: In the field of supply chain management, what kinds of big data analytic models are being utilized?	The big data analytic models that are utilized for the supply chain should be categorized.	Big data models
RQ4: In order to create these models, what kinds of big data analytics approaches are utilized?	Make a list of the primary methods of big data analysis that are utilized in the models.	Big data techniques
RQ5: What kinds of search strategies are being utilized in the works?	Make a list of the most common search strategies.	Search methods
RQ6: The studies were published in which locations?	Provide a list of the vehicles that were utilized for disclosing the results.	Search location

There is no doubt that business data analysis (BDA) is being utilized in the management and control areas of the chain. This is due to the fact that the incorporation of information into decision-making processes confers a significant competitive advantage. The production sectors of the chain, on the other hand, stand to benefit greatly from the deployment of BDA because it allows for the optimization of a great number of processes, which ultimately results in an increase in the value of the product. Big data analytics can be categorized into various tiers. To better comprehend the application of BDA principles, this research topic aims to identify the most common levels of application.

An explanation of the primary findings, which are presented in Table 3, is provided in the following sections. Among the levels that were discovered, the levels of predictive and descriptive analysis are represented by 66 percent, or 37 out of the fifty papers that were evaluated. This is significant because the levels were discovered. One of these distinct types of study is known as the predictive analysis, and its primary objective is to identify potential future moves by utilizing the data that is being supplied by the most varied platforms. On the other hand, descriptive analysis strives to discover links and/or associations between historical data, and it forecasts future movements based on these facts with the intention of predicting future movements. There is the potential for comparisons to be made between the two distinct types of analysis. The levels of mixed and prescriptive BDA account for 34 percent of the fifty articles that were evaluated, which is equivalent to 17 of the articles. This percentage is based on the analysis of individual publications. The prescriptive level is concerned with the tools and/or methods that are utilized for the analysis and presentation of information acquired by BDA approaches, whereas the mixed level is concerned with the degree of BDA that employs more than one approach to achieve the results. The mixed level is concerned with the degree to which BDA employs more than one approach to achieve the results. Both of these levels are relatively generic in nature.

Table 3: Survey questions

Id	Question
Q1	Taking into consideration the procedures involved in the supply chain, do you think that big data analytics approaches are crucial for management?
Q2	When it comes to the processes involved in supply chain management, do you believe that big data analytics can help with managing goods while it is in transit?
Q3	Regarding the procedures involved in the supply chain, do you agree that big data analytics can be of assistance in the process of vehicle routing (logistics)?
Q4	Do you agree that big data analytics can help in identifying selection for installations, taking into consideration the procedures involved in supply chain management?
Q5	Taking into consideration the procedures involved in the supply chain, do you agree that big data analytics can help in the selection of suppliers?
Q6	When taking into consideration the procedures involved in supply chain management, do you agree that big data analytics can help with demand-driven storage?
Q7	Do you agree that big data analytics can be of assistance in real-time demand procedures, taking into

	consideration the processes involved in supply chain management?
Q8	When it comes to supply chain procedures, do you think big data analytics can assist cut costs?
Q9	Will there be an abundance of merchandise in the gondola? Isn't big data analytics a useful tool for supply chain processes?
Q10	To what extent do you think big data analytics may improve order collection processes inside the supply chain? You brought up some good issues about the supply chain in your earlier questions. Pick three processes—the ones that, in your opinion, are most crucial.
Q11	Increase your influence all the way through

Technology and our relationship with it have been evolving over the years, bringing with it an exponential growth in the amount of data produced by a wide variety of sources. As a result, algorithms have been developed to sift through this deluge of data, find patterns, and make predictions regarding individual cases. RQ 3: Models for Analyzing Big Data. Focusing on the most popular BDA models, this research query returns them. For a total of 48 percent, or 24 out of 50 articles, the optimization and prediction models were considered. It is the goal of the optimization model to use data in an effort to improve some part of the chain, be it the time spent on it, the money saved, or the efficiency of the processes involved. Forecasting, on the other hand, is more commonly linked with management and decision-making, and its primary purpose is to produce predictions for improved positioning. Of the fifty articles that were reviewed, sixteen (42% of the total) were related to classification and simulation models.

In order to concentrate on the activities that are most important, the classification model provides a list of the chain's primary processes.

The simulation processes, on the other hand, concentrate on replicating future situations. The ability to foresee and avert problems makes these procedures ideal for use in production. Out of fifty articles that were studied, ten (or twenty percent) fall into the other models, which include visible, mixed, and others. All of the models are interconnected because they include data visualization solutions for information management. They may be the less prominent models, but they are related models nonetheless. They play an important function, as demonstrated in earlier sections; yet, they have not been extensively researched using a BDA model. This is because our understanding of them is still limited. By looking at the most popular BDA models in the papers, we can see how the optimization and prediction models are represented more thoroughly. Information management and decision-support have been the BDA's primary areas of focus, in contrast. While these models can certainly serve this function, it would be intriguing to go into the reasons behind the underutilization of visibility reasons.

RQ4: Big Data Analytic Techniques. Specifically, this research query brings up the BDA strategies that are utilized the most commonly. 64 percent, or 34 out of the fifty publications that were reviewed, are comprised of mixed approaches, visualization, and heuristic approach. Due to the fact that the mixed approach, as its name suggests, employs two or more techniques in order to purchase a BDA structure, these techniques have a highly representative representation. Data mining techniques, which have become increasingly popular over the past few years, are typically employed in conjunction with the visualization technique, which is one of the strategies that can be found in the majority of articles that have mixed classification. It is acceptable to apply this method on a big scale given that the heuristic approach is one of the most extensively used BDA models. In contrast, optimization using a heuristic is commonplace.

The other methods that were presented account for 32 percent, or sixteen out of fifty articles. These are methods that, despite their significance, are associated with extremely particular processes and, as a result, are utilized in lower amounts; they are typically utilized in conjunction with other methods.

In addition to being strategies that are centered on giving information and optimizing processes, it is possible to observe a trend to employ techniques that we may refer to as generic. This is something that can be observed because these approaches can be implemented at all stages of the chain. On the other hand, particular methods have been utilized in a limited capacity, despite the fact that they offer a significant difference.

RQ5: Research Models. This research question aims to gain an understanding of the research methodologies that were utilized by the studies that were chosen. According to [16], there are six different types of research methodologies that can be used to classify studies. Research on solution proposals, evaluations, validations, opinions, experiences, and philosophical articles fall into this category. Of the fifty articles, 66% are from evaluation research models and 4% are from solution research models. These models suggest that the articles are being published to assess the current chain processes and/or to suggest new ways to improve them. Only 34% of the 50 papers included in the analysis are in the other models. Opinion models and the verification of research concepts and techniques are two examples of their contributions.

At last, it's clear that the majority of the published works in the fields under consideration aim to either evaluate the present level of knowledge in terms of methodologies and ideas or provide solutions to issues that crop up

throughout the processes. Because modern technologies provide so many opportunities to enhance processes, this kind of behavior is becoming more common as a result of them.

RQ6: Research Location. The publication dates and venues of the papers under consideration are the focus of this research topic. The primary goal is to take a look at how supply chain activities are being analyzed utilizing big data analytics in current studies. That is why we made note of the publication year and source for every single study that was chosen. All of the relevant studies' timelines are shown in Figure 3. From 2010 to 2021, Figure 3 shows a solid gray line that represents the number of evaluated research publications released each year.

No studies were located before to 2010, even though the search approach only covers the period till 2021. This might have happened since the initial definition of "big data" was in 2001, when the concept was far from what it is now; it might have taken some time until people really grasped and could make good use of these new technologies. There was a dramatic uptick in studies examining BDA in the supply chain from 2011 to 2014, but 2016 saw the greatest concentration of studies due to the big data explosion.

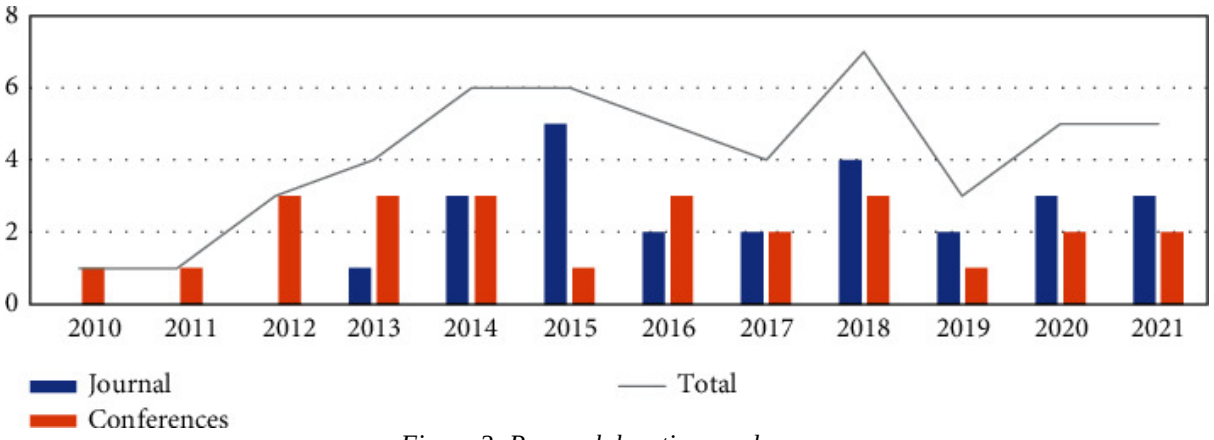

Figure 3: Research locations and year

During the period of time spanning from 2016 to 2021, it is evident that there was a substantial rise in the quantity of research projects that were focused on the topic area. With that being said, it will be necessary to keep a close eye on the years that follow in order to verify that BDA has been extensively deployed for activities related to supply chain management ever since the year 2021, when the polls had a slight decrease. This will allow for the confirmation of the aforementioned statement.

Conclusion

The major goal of this research was to categorize existing literature and then to identify key sectors of the supply chain that utilize big data analytical methods. It was hoped that both of these goals would be achieved. In order to create a connection between the numerous issues that are currently being explored, the academic community came up with the notion of conducting a survey and creating a comparison sample. A total of fifty articles were chosen for the purpose of analysis, and the filtering method that was utilized for the mapping study was utilized. After that, the research questions were used to classify these papers into different categories. With the use of these research questions, it was easy to determine the path that the study will go. Over seventy percent of the BDA principles were found to be applied in the administration of some process within the supply chain, according to the findings of the project. This procedure required a significant amount of time, the vast majority of which was spent on controlling it, while the remaining time was spent ensuring that demand was satisfied.

References

[1]. Chong D., Shi H. Big data analytics: a literature review. Journal of Management Analytics. 2015;2(3):175–201. doi: 10.1080/23270012.2015.1082449. [CrossRef] [Google Scholar]

[2]. Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. Production and Operations Management, 27(10), 1868–1883.

[3]. Majidian S., Vanani R., Iman. Literature Review on Big Data Analytics Methods. London, UK: Intechopen; 2019.

[4]. Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. Future Generation Computer Systems, 101, 993–1004.

[5]. Potočnik, P., Šilc, J., Papa, G., et al. (2019). A comparison of models for forecasting the residential natural gas demand of an urban area. Energy, 167, 511–522.
[6]. Ning, C., & You, F. (2019). Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. Computers and Chemical Engineering, 125, 434–448.
[7]. Mathrani S., Lai X. Big data analytic framework for organizational leverage. Applied Sciences. 2021;11(5):p. 2340. doi: 10.3390/app11052340. [CrossRef] [Google Scholar]
[8]. Solanki M. S., Sharma M. A. A literature review on big data analytics. International Journal of Innovative Research in Computer Science & Technology. 2021:234–247. doi: 10.55524/ijircst.2021.9.6.54. [CrossRef] [Google Scholar]
[9]. Alsaffar M., Hamad A. A., Alshammari A., et al. Network management system for IoT based on dynamic systems. Computational and Mathematical Methods in Medicine. 2021;2021:8. doi: 10.1155/2021/9102095.9102095 [PMC free article] [PubMed] [CrossRef] [Google Scholar] Retracted
[10]. Hamad A. A., Thivagar M. L., Alshudukhi J., et al. Secure complex systems: a dynamic model in the synchronization. Computational Intelligence and Neuroscience. 2021;2021:6. doi: 10.1155/2021/9719413.9719413
[11]. Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and Industry 4.0 applications.
[12]. Alhameli, F., Ahmadian, A., & Elkamel, A. (2021). Multiscale decision-making for enterprise-wide operations incorporating clustering of high-dimensional attributes and big data analytics: Applications to energy hub. Energies, 14(20).
[13]. Niu, B., Dai, Z., & Chen, L. (2022). Information leakage in a cross-border logistics supply chain considering demand uncertainty and signal inference. Annals of Operations Research, 309(2), 785–816.
[14]. Sengan S., Khalaf O. I., Priyadarsini S., Sharma D. K., Amarendra K., Hamad A. A. Smart healthcare security device on medical IoT using raspberry pi. International Journal of Reliable and Quality E-Healthcare. 2022;11(3):1–11. doi: 10.4018/ijrqeh.289177.
[15]. Dolgui, A., & Ivanov, D. (2022). 5G in digital supply chain and operations management: Fostering flexibility, end-to-end connectivity and real-time visibility through internet-of-everything. International Journal of Production Research, 60(2), 442–451.
[16]. Li, L., Gong, Y., Wang, Z., & Liu, S. (2022b). Big data and big disaster: A mechanism of supply chain risk management in global logistics industry. International Journal of Operations and Production Management.