International Journal of Information Management 49 (2019) 86–97

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing

Ian M. Cavalcante[a], Enzo M. Frazzon[b], Fernando A. Forcellini[a], Dmitry Ivanov[c,*]

[a] Mechanical Engineering Department, Federal University of Santa Catarina, Florianópolis, Brazil
[b] Production and Systems Engineering Department, Federal University of Santa Catarina, Florianópolis, Brazil
[c] Supply Chain Management, Berlin School of Economics and Law, Berlin, Germany

ARTICLE INFO

Keywords:
Supplier selection
Machine learning
Simulation
Digital supply chain
Data-driven decision-making support
Resilience
Digital supply chain twin

ABSTRACT

There has been an increased interest in resilient supplier selection in recent years, much of it focusing on forecasting the disruption probabilities. We conceptualize an entirely different approach to analyzing the risk profiles of supplier performance under uncertainty by utilizing the data analytics capabilities in digital manufacturing. Digital manufacturing peculiarly challenge the supplier selection by the dynamic order allocations, and opens new opportunities to exploit the digital data to improve sourcing decisions. We develop a hybrid technique, combining simulation and machine learning and examine its applications to data-driven decision-making support in resilient supplier selection. We consider on-time delivery as an indicator for supplier reliability, and explore the conditions surrounding the formation of resilient supply performance profiles. We theorize the notions of risk profile of supplier performance and resilient supply chain performance. We show that the associations of the deviations from the resilient supply chain performance profile with the risk profiles of supplier performance can be efficiently deciphered by our approach. The results suggest that a combination of supervised machine learning and simulation, if utilized properly, improves the delivery reliability. Our approach can also be of value when analyzing the supplier base and uncovering the critical suppliers, or combinations of suppliers the disruption of which result in the adverse performance decreases. The results of this study advance our understanding about how and when machine learning and simulation can be combined to create digital supply chain twins, and through these twins improve resilience. The proposed data-driven decision-making model for resilient supplier selection can be further exploited for design of risk mitigation strategies in supply chain disruption management models, re-designing the supplier base or investing in most important and risky suppliers.

1. Introduction

Companies whose suppliers are prone to disruption risks have a common question to ask. How do firms obtain better performance than others if similar suppliers are affected by disruptions? Recent research hypothesized that some of that success is attributable to the resilient supplier selection and development (Gao, Simchi-Levi, Teo, & Yan, 2018; Hosseini & Barker, 2016; Hosseini et al., 2019b; Kull & Talluri, 2008; Narasimhan & Talluri, 2009; Sawik, 2013a; Torabi, Baghersad, & Mansouri, 2015; Yoon, Talluri, Yildiz, & Ho, 2018). Manufacturing firms operate in environments with inherent uncertainties in demand, supply, cost, lead time (LT) and catastrophic disasters (Ivanov, Dolgui, & Sokolov, 2019; Papadopoulos et al., 2017; Rajagopal, Venkatesan, & Goh, 2017). The increase in data availability and the emergence of new digital technologies, such as machine learning, cloud computing, internet of things (IoT) and blockchain enable managers and government to cope with uncertainties using intelligent decision-making principles (Dubey, Gunasekaran, & Childe, 2015; Grant & Yeo, 2018; Gunasekaran, Kumar Tiwari, Dubey, & Fosso Wamba, 2016; Ismagilova, Hughes, Dwivedi, & Raman, 2019; Kshetri, 2018; Kumar, Mangla, Luthra, Rana, & Dwivedi, 2018; Liu, Chan, Yang, & Niu, 2018; Rana et al., 2018). The Big Data phenomenon forced the development of new techniques in fast analytics and data science as part of business intelligence using firms' dynamic capabilities (Akter et al., 2019; Duan, Edwards, & Dwivedi, 2019; Larson & Chang, 2016; Wamba et al., 2017). Altay, Gunasekaran, Dubey, and Childe (2018) point out supply chain agility and supply chain resilience are dynamic capabilities that have significant effect on supply chain performance.

Digital manufacturing peculiarly challenge the supplier selection by the dynamic order allocations, and opens new opportunities to exploit

* Corresponding author.
E-mail address: dmitry.ivanov@hwr-berlin.de (D. Ivanov).

https://doi.org/10.1016/j.ijinfomgt.2019.03.004
Received 2 February 2019; Received in revised form 11 March 2019; Accepted 11 March 2019
Available online 28 March 2019
0268-4012/

the digital data to improve sourcing decisions. Gandomi and Haider (2015) believe the current hype can be attributed to leading technology companies, such as IBM, who invested in building a niche analytics market. Techniques involving supervised machine learning (SML) have already become powerful tools with various applications within intelligent manufacturing systems (Jordan & Mitchell, 2015; Wuest, Weimer, Irgens, & Thoben, 2016). In this study, SML is further investigated regards to its application to supplier selection in digital manufacturing with consideration of resilience.

Supplier selection is a critical issue for maintaining competitive advantage in supply chain (SC) management (Dey, Bhattacharya, Ho, & Clegg, 2015; Dickson, 1966; Weber & Current, 1993; Wetzstein, Hartmann, Benton, & Hohenstein, 2016). Multi-factor supplier selection has been recently extended by inclusion of disruption risks to address SC resilience (Ang, Iancu, & Swinney, 2016; Das, Narasimhan, & Talluri, 2006; Gao et al., 2018; He, Alavifard, Ivanov, & Jahani, 2018; Hosseini & Barker, 2016; Ivanov, Dolgui, Sokolov, & Ivanova, 2017; Tomlin, 2006; Wu & Olson, 2008; Yoon et al., 2018). Achieving SC resilience involves adopting reactive and proactive approaches by creating certain protections and taking into account possible perturbations through contingency plans or backup supply planning (Elluru, Gupta, Kaur, & Singh, 2017; Hosseini et al., 2019a). Digital technologies do not only enable data-driven decision support tools (Frazzon, Kück, & Freitag, 2018), but also stimulate the development of new production forms, such as smart manufacturing and Industry 4.0 (Ivanov et al., 2019; Kumar, Singh, & Lamba, 2018; Kusiak, 2018; Liao, Deschamps, Loures, & Ramos, 2017; Rossit, Tohmé, & Frutos, 2018). These new forms of digital manufacturing are characterized by higher flexibility, make-to-order environments and customer-driven SC dynamic structuring, which requires dynamic supplier selection analysis. At the same time, digital manufacturing is expected to face increased disruption risks due to increasing complexity and globalization (Ivanov et al., 2019). As such, there is also a need for new modeling approaches with which to analyze resilient supplier selection in novel organizational networks (Durugbo, Tiwari, & Alcock, 2013) and Big Data can be essential in supplier risk management as it can give detailed understanding of supplier performance towards the identification of opportunities for better sourcing (Lamba & Singh, 2017).

Despite the considerable progress in resilient supplier selection and data-driven decision support systems in SCs, we are not aware of any published research that considers data-driven approaches to supplier selection with consideration of resilience in a digital manufacturing environment. Since digital data is considered a key source for both new manufacturing forms and decision-support systems (Chae, 2019), the objective of this study is to close the research gap described above and, by means of a test case, advance knowledge of how SML can contribute to supplier selection in the context of digitalization and SC resilience. Supplier selection involves consideration of both recurrent and disruption events, i.e., frequent events with low impact and rare events with high impact, respectively (Dolgui, Ivanov, & Sokolov, 2018; Ivanov, 2018; Tomlin, 2006; Yoon et al., 2018). According to Sawik (2016), in make-to-order environments and customer-driven SCs, customer service level is of particular importance, since it can be analyzed as a substitute for shortage costs that are hard to estimate.

While studies have established a salience of resilient supplier selection in recent years, much of it was focusing on forecasting the disruption probabilities. We conceptualize an entirely different approach to analyzing the risk profiles of supplier performance under uncertainty by utilizing the data analytics capabilities in smart manufacturing using the digital twin principles. Hence, uncertainties of the environment are evaluated based on the learning of the system in terms of suppliers performance in the SC. Digital SC twins were recently defined as computerized models that represent the network state for any given moment in time and allow for complete end-to-end SC visibility to improve resilience and test contingency plans (Ivanov, 2019).

We develop a hybrid technique, combining simulation and SML and examine its applications to data-driven decision-making support in resilient supplier selection. We consider on-time delivery (OTD), also known as delivery reliability, an indicator for supplier reliability, and explore the conditions surrounding the formation of resilient supply performance profiles. Further, we theorize the notions of risk profile of supplier performance and resilient SC performance. We define risk profile of supplier performance as a set of negative outcomes in simulation runs associated with a particular supplier in terms of failing to meet the OTD requirement regarding the orders allocated at the supplier. Resilient supply chain performance, in turn is abstracted to total OTD of the SC regarding the customer demand (Ivanov, Tsipoulanidis, & Schönberger, 2017; Karim, Samaranayake, Smith, & Halgamuge, 2010; Wieland & Marcus Wallenburg, 2013) as a composition of delivery date and delivery quantity.

We show that the associations of the deviations from the resilient supply chain performance profile with the risk profiles of supplier performance can be efficiently deciphered by our approach. The results suggest that a combination of SML and simulation, if utilized properly, improve the delivery reliability. Such a combination is unique in literature. It mimics the complexity of business reality affording a more realistic approach to making sourcing decisions and appears to be more relevant in practical environments. Our approach can also be of value when analyzing the supplier base and uncovering the critical suppliers, or combinations of suppliers the disruption at which results in the adverse performance decreases. The results of this study contribute to the understanding of the use of digital SC twins with the aim to improve resilience by means of the combination of machine learning (ML) and simulation. The proposed resilient supplier selection can be further considered as a support system to design risk mitigation strategies in the development of supplier portfolio, SC disruption management models or investing in most important and risky suppliers.

The remainder of this paper is organized as follows. Section 2 presents current relevant literature concerning supplier selection approaches and data-driven decision support systems with consideration of resilience. Section 3 describes a hybrid approach to supplier selection, combining simulation and ML. The results are discussed in Section 4 and managerial and theoretical implications are discussed in Section 5. Section 6 concludes the paper by highlighting the insights gained in this study and outlining future research opportunities.

2. State-of-the-art

2.1. Supplier selection with resilience considerations

Our study builds on three conceptual perspectives. First and principally, we greatly benefited from the literature on resilient supplier selection. The second perspective is the application of ML techniques to SC management. Finally, studies on data-driven decision support systems for supplier and disruption risk management pointed the directions in development of information management framework in our model.

As SC structures become increasingly complex and global, manufacturing firms become increasingly dependent on their suppliers. Complexity and globalization also increase SC risk exposure (Gao et al., 2018). Hamdi, Ghorbel, Masmoudi, and Dupont (2018) state that the best supplier is usually the one who can deliver the right product at the right time, in the right place, in the right quantity at a competitive price. Increased SC risk exposure forces the inclusion of resilience into supplier selection procedures (Gao et al., 2018; He et al., 2018; Hosseini & Barker, 2016; Ivanov, Dolgui, et al., 2017; Larson & Chang, 2016; Weber & Current, 1993). The concept of SC resilience is defined by Ivanov et al. (2019) as a complex characteristic of non-failure operation, durability, recoverability with the maintenance of SC processes and the SC as a whole. Therefore, supplier selection has become a key element in designing efficient, synchronized and resilient operations in digital manufacturing SCs.

Rajagopal et al. (2017) and Hamdi et al. (2018) carry out literature reviews correlating topics of supplier selection and risk management in the SC, the first being a subset of the second. According to Tomlin (2006) and Dolgui, Ivanov, et al. (2018), the risks in SCs can be divided into recurrent or disruptive. For Jüttner, Peck, and Christopher (2003), risks in the SC can be classified as internal risks, SC risks or external risks. The authors state the first arises within the organization, the second arises externally to the organization but within the SC, and the third arises externally and outside the SC, that is, it affects several chains simultaneously.

Ivanov and Dolgui (2018) elaborate on the importance of increasing SC resilience in efficient ways, i.e., achieving SC resileanness (resilient + lean). As such, resilient supplier selection must be subject to a multi-factor analysis. In his seminal paper, Dickson (1966) presents 23 criteria, of which quality, delivery and performance history are shown to be the three most important factors in vendor selection. Drawing from recent literature, Chen, Hsieh, and Wee (2016) cluster 11 main criteria for supplier selection, alongside their definitions, these are: finance, quality, delivery, relationship, service, technology, supply facilities, management, efficacy, environment and risk factors. Therefore, in this paper, delivery reliability assessed based on historical data is considered to measure service level performance. Furthermore, Hamdi et al. (2018) subdivides supplier selection decision approaches into four: (i) quantitative, (ii) qualitative, (iii) using simulation tools and (iv) using artificial intelligence. In quantitative approaches, the factors or criteria under study can be measured and quantified numerically, for instance, delivery reliability (Wetzstein et al., 2016). None of these studies, though, formally examined the data analytics capabilities in selecting the resilient supplier portfolios – a distinctive and significant contribution made by our study. In our study, a hybrid approach integrating a simulation tool and SML techniques is developed and tested.

Rajesh and Ravi (2015) state resilience means the adaptive capability to respond to disruptions and recover from them. A resilient supplier is able to provide good quality products at economic rates and is flexible enough to accommodate demand fluctuations with shorter LTs over a lower ambience of risk without compromising safety and environmental practices (Rajesh & Ravi, 2015). Furthermore, the development of SC resilience is of particular importance when developing an agile SC in uncertain market conditions and flexibility strategies. For example, dual or multiple sourcing are typically utilized to cope with disruption risks and recovery measures (Gosling, Purvis, & Naim, 2010; He et al., 2018; Ivanov, Dolgui, et al., 2017).

Torabi et al. (2015) propose a five-step method to enhance the supply resilience level in a scenario-based, bi-objective, possibilistic mixed integer linear model to build resilient supply bases for global SCs. The computational experiments indicate the consideration of disruptive events can have significant impact on selected supply bases. The authors introduce a new supply side objective function to calculate the resilience level of the selected supply base and consider several strategies, such as suppliers' business continuity plans, fortification of suppliers and contracts with backup suppliers, to enhance the resilience level of the supply network.

Sawik (2013a) states a resilient supply portfolio includes protected suppliers that are capable of supplying despite disruption, as well as having emergency inventory options which can be used to compensate for the lost capacity of suppliers and to replace non-delivered parts ordered from disrupted suppliers.

Bohner and Minner (2017) use a mixed-integer linear program approach to solve the supplier selection problem subject to disruptions. The model considers backup suppliers and less risky, but more expensive, main supplier. They evaluate supplier selection performance in terms of cost and trade-offs between economies of scale and failure risk. Diverse techniques and methods, such as multi-objective mixed integer programming, stochastic mixed integer programming, fuzzy analytic hierarchy process, Bayesian network, conditional value-at-risk (CVaR), worst-case CVaR, data envelopment analysis, technique for order of preference by similarity to ideal solution and analytical approaches are used to perform resilient supplier selection under operational and disruption risks (Ang et al., 2016; Chen, 2011; Dupont, Bernard, Hamdi, & Masmoudi, 2018; Gao et al., 2018; PrasannaVenkatesan & Goh, 2016; Sawik, 2013b; Viswanadham & Samvedi, 2013).

Vugrin, Warren, and Ehlen (2011) define the resilience capacity of a system as a function of absorptive, adaptive, and restorative capacities. Hosseini and Barker (2016) used this concept and extended it to the supplier selection problem. First, absorptive capacity refers to the ability to absorb shocks from disruptive events, implying proactive planning or development of pre-disaster strategies. In the context of supplier selection, the authors cite four main features: geographical segregation, i.e., segregation of a supplier geographically from natural disasters, surplus inventory, backup supplier contracting and physical protection, i.e., security of suppliers' facility from disruptive events. Second, adaptive capacity is considered a temporary post-disaster strategy, e.g., redundant transportation for use in non-standard rerouting following disruption. Last, restorative capacity refers to the recover phase and is the last line of defense against disruption.

Considering the SC as a whole, the dynamic nature of the supplier-customer relationship influences disruption propagation, and therefore the SC structure and dependence (Scheibe & Blackhurst, 2018). Wu and Olson (2008) state long-term and permanent relationships in SCs usually result in benefits such as lower purchase costs and can culminate in lower prices for the final customer. This perspective is echoed by Sheffi and Rice (2005) and Chen et al. (2016), who emphasize that loyalty in the supplier-customer relationship provides benefits to the SC by making it more resilient to crisis and demand fluctuations. However, the authors do not address the fact that in this period of digital transformation firms are becoming more data-oriented and may even overlap this loyal relationship between supplier-customer in decision-making supplier selection processes.

Despite the significant advances achieved in recent years, the literature reviewed does not specify an explicit approach to using the digital data in improving SC performance by building the resilient supplier portfolios. While a growing body of literature pointed to the importance of developing the resilient SC, less attention was directed to the exploiting the resilience capabilities through a dynamic analysis of SC performance (Ivanov, 2018; Ivanov, Tsipoulanidis, et al., 2017). An important dimension in resilient supplier selection – the dynamic analysis of supplier performance risk profiles was left ignored. Given that the relationship between suppliers and customers may become ephemeral and strongly influenced by data with automated intelligent decision-making, it is possible to perceive that new research opportunities in this field will arise.

2.2. Data-driven decision support systems for supplier and disruption risk management

Digital factory concepts share the attributes of smart networking (Strozzi, Colicchia, Creazza, & Noè, 2017). The vision of Industry 4.0 is that the manufacturing system contains all the relevant information about its production and supply requirements. Digital technologies enable flexible decision-making by providing real-time data for all parts of the SC (Bounfour, 2016; Dubey et al., 2019; Wamba et al., 2017).

Dubey et al. (2015), Papadopoulos et al. (2017), Gunasekaran et al. (2016), Choi, Wallace, and Wang (2018) and Nguyen, Li, Spiegler, Ieromonachou, and Lin (2018) provide evidence that data analytics is being applied to SC management in procurement, manufacturing shop floors, routing optimization, real-time traffic operation monitoring, proactive safety management, and in-transit inventory management in logistics/transportation. Reducing SC cost as well as carbon emissions are important tasks to consider in operational decisions in order to be competitive in the digital manufacturing environment (Lamba & Singh, 2018; Lamba, Singh, & Mishra, 2019). Models providing optimal decisions considering sustainable procurement and transportation based on real data can be found in the literature (Kaur & Singh, 2017, 2018). Furthermore, Kaur and Singh (2016) model sustainability-resilience link at the supply chain design level through the procurement and logistics of raw material. Their model suggests there is a trade-off between lot-size orders, carbon emissions and SC resilience, meaning that smaller lot-size leads to larger carbon emission due to transportation and greater risk of supply chain disruption. A similar problem setting of sustainable use of resources to build SC resilience can be found in Pavlov, Ivanov, Pavlov, and Slinko (2019).

Papadopoulos et al. (2017) point out that data analytics can help improve SC risk management and disaster-resistance. Choi and Lambert (2017) and Choi, Chan, and Yue (2017) provide evidence of how data analytics can be used to improve the resilience of SC operations by utilizing firms databases and large volumes of data to predict risks, assess vulnerability and enhance their SCs.

Simchi-Levi et al. (2015) present a data-driven system to analyze supplier exposure in the automotive sector. This system estimates supplier risk exposure, and evaluates pre-disruption risk mitigation actions and optimal post-disruption contingency plan deployment. The system integrates databases, a quantitative risk-exposure model, and an output performance visualization tool. The data sources include material requirements planning system, the purchasing database, and sales-volume planning information based on the SC mapping methodology (Gusikhin & Klampfl, 2012). The optimization engine uses the data to test the various performance impacts of disruptions. Decision-makers in procurement and risk specialists can use the system to track risk exposures in real time as inventory levels fluctuate and the SC structure evolves. The frequency of updates relies on the data-integration technology and the computational tractability of the optimization models.

Ivanov et al. (2018) show that data analytics can be used at the planning stage to identify supplier risk exposure and can help at the reactive stage to monitor and identify disruptions. They propose a framework of integrated cyber-physical SC simulation and optimization and relate this framework to system-cybernetics principles. Their results echo those in the study by Choi (2018) that presented a new practical perspective on how big data related technologies can be used for global SCs with a system of systems mindset.

2.3. ML applications to SCs and manufacturing

ML can be applied to resilient SCs. Baryannis, Validi, Dani, and Antoniou (2018) summarize recent AI applications to SC risk management and show future research opportunities in risk identification, assessment and response. Priore, Ponte, Puente, and Gómez (2018) apply ML to the dynamic selection of replenishment policies according to SC environmental dynamics. ML techniques have been applied to detect bottlenecks, high-risk tasks and events in order to achieve adequate production rescheduling (Dolgui, Bakhtadze, et al., 2018; Ji & Wang, 2017). Palombarini and Martínez (2012) prototype an application that performs rescheduling based on relational reinforcement learning (RL).

Shahzad and Mebarki (2012) propose framework based on data mining for job shop scheduling problems (JSSPs) that identifies the critical parameters and states of particular dynamic scheduling environments. Stricker, Kuhnle, Sturm, and Friess (2018), Waschneck et al. (2018) and Li, Wang, and Sawhney (2012) use RL to solve the JSSP. First, Stricker et al. (2018) develop an RL-based adaptive order dispatching algorithm that can outperform existing rule-based heuristics approaches. Second, Waschneck et al. (2018) test an RL approach in a simulation of a discrete event at a small semi-conductor factor and observe that although the learning algorithms do not overcome the heuristics, the RL was able to reach an expert knowledge level with two days of training. Third, Li et al. (2012) investigate pricing, lead-time, scheduling and order acceptance decisions in a make-to-order manufacturing system with stochastic demands in a discrete-event simulation model. They develop an RL based Q-learning algorithm (QLA) and find that the QLA performance is superior to the existing policies.

Tuncel, Zeid, and Kamarthi (2014) apply an RL approach to solve a disassembly line balancing problem with uncertainty. Kartal, Oztekin, Gunasekaran, and Cebi (2016) develop a hybrid methodology that integrates ML with multi-criteria decision-making techniques in order to execute multi-attribute inventory analysis. The authors implemented naive Bayes, Bayesian network, artificial neural networks (NN), and support vector machine (SVM) algorithms to predict classes of initially determined stock items in a large-scale automotive company. Sharp, Ak, and Hedberg (2018) analyze approximately 4000 abstracts by means of the Natural Language Processing technique and conclude that generically applicable algorithms such as NNs and SVMs are gaining popularity in the field of manufacturing.

Another application of ML to manufacturing is prediction of LT and cycle time (CT) key performance indicators. Most production planning and scheduling methods rely on LTs. The efficiency of these methods is crucially affected by the accuracy of LT prediction (Gyulai et al., 2018). The authors perform an LT prediction based on regression algorithms for a real flow-shop environment exposed to frequent changes and uncertainties resulting from the changing customer order stream. Lingitz et al. (2018) use SML approaches to perform LT prediction based on historical production data obtained from manufacturing execution systems. CT forecasting is one of the most crucial issues for production planning in terms of maintaining high delivery reliability in semiconductor wafer fabrication systems (Wang, Yang, Zhang, Wang, & Zhang, 2018). Wang, Zhang, and Wang (2018) use a recurrent NN to model a CT forecast, estimating the short-term CT forecast of wafer lots.

Location awareness has high potential to produce valuable information in manufacturing facilities (Carrasco, Coronado, Parto, & Kurfess, 2018). Technologies such as radio frequency identification (RFID) and bluetooth low energy devices, e.g., beacons, enable the collection of data pools from manufacturing shop-floors. Carrasco et al. (2018) present a system that finds the nearest machine to a user. The authors use nearest neighbor, weighted k-Nearest Neighbor (k-NN) and Bayesian inference techniques. Solti, Raffel, Romagnoli, and Mendling (2018) investigate the effectiveness and efficiency of outlier detection methods for finding misplaced products in a real setting with an RFID inventory robot. Their research suggests that ML techniques can be effectively used to harness sensor systems for improved operational use cases. Similarly, Kho, Lee, and Zhong (2018) use RFID technology to capture real-time production data and then apply two ML techniques: k-means clustering and gradient descent optimization. The authors state that valid predictions about the expected overall manufacturing time for a given number of manufacturing batch inputs can be obtained.

ML has been used to improve manufacturing at the process level. For instance, Diaz-Rozo, Bielza, and Larra naga (2017) propose a cyber-physical system (CPS) for machine component knowledge discovery based on clustering algorithms using real data from a machining process. Three clustering algorithms are compared – k-means, hierarchical agglomerative and Gaussian mixture models – in terms of their contribution to spindle performance knowledge during high throughput machining operation. Furthermore, Kruger, Shih, Hattingh, and van Niekerk (2011) show that the process optimization is capable of learning and optimizing a high-volume gun drilling process. The learning process generated regression models for the manufacturing process and the agent was able to determine the optimal trade-off between the technical and economic factors.

Guo, Yuan, and Tian (2009) present a SVM model combined with decision tree (DT) to address issues on supplier selection including feature selection, and multi-class classification. Mirkouei and Haapala (2014) also use SVM and DT integrated with a mathematical programming approach to supplement existing supplier selection methods in a biomass-to-biofuel SC.

Although ML is not a favorable method for all industrial problems, encouraging the application of learning algorithms can contribute to the achievement of autonomous production systems (Stricker et al., 2018). Kusiak (2017) highlights five gaps in manufacturing innovation in the digital transformation era: (i) adopt data strategies, (ii) improve data collection, use and sharing, (iii) design predictive models, (iv) study general predictive models and (v) connect factories and control processes. Therefore, since ML provides intelligent outcomes from data, a close follow up in this research field is fundamental to innovation in a resilient data-driven manufacturing environment.

Despite of significant advances in ML application to SC and operations management achieved recently, the literature does not specify directions as to how to make use of digital data and to utilize the ML advantages to build resilient supplier portfolios. As a result, it is not yet clear how ML can contribute to the conceptual and technological frameworks of resilient supplier selection. This also means that the causes of SC performance perturbations due to disruptions in supply base have not been entirely disentangled from the risk profiles of supplier performance.

3. Digital manufacturing experimentation

In this section, a digital manufacturing experiment is described, which adopts a hybrid approach in combination with simulation and ML models and integrates these within the context of supplier selection.

3.1. Simulation model

Simulations make use of agents, system dynamics and discrete events to gain a better understanding of interactions and support the deployment of organizational networks (Durugbo et al., 2013). In this study, the simulation model is performed with Anylogic software and represents a make-to-order manufacturing system which has up to four raw material suppliers. Fig. 1 illustrates the information and materials flows in the simulation model.

According to the model parameters, raw material orders only occur after a customer order is consolidated and raw material is the only necessary supply to manufacture the final product. The purchase orders are characterized by normal distributions as shown in Bodaghi, Jolai, and Rabbani (2018) as well as demand uncertainty. Furthermore, it is assumed only one type of product being delivered and price and supplier competition analysis are neglected.

Supplier performance is modeled in way that is similar to that of Tomlin (2006): one supplier may be unreliable in a certain period and also may have deterministic capacity limitations. In this paper, four possible suppliers are considered and the previously mentioned restrictions influence the delivery performance of suppliers, which is modeled according to a normal distribution.

3.2. ML model

ML addresses the question of how to build computers that improve automatically through experience. It is one of the most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the edge of artificial intelligence and data science (Jordan & Mitchell, 2015). In this work, the ML model is implemented using the Scikit-Learn package, which is defined by Pedregosa et al. (2011) as a Python module that integrates a wide range of state-of-the-art ML algorithms for medium-scale supervised and unsupervised problems. Moreover, other packages such as Numpy, Matplotlib and Pandas are also used to perform data preprocessing, data analysis and visualization tasks. Fig. 2 shows the supplier selection model using SML.

The preprocessing step can often have a significant impact on the generalized performance of a SML algorithm and may include substeps, such as data cleaning, normalization, transformation, feature extraction and selection, etc. (Kotsiantis, Kanellopoulos, & Pintelas, 2006). In this work, since data is generated from a simulation, the database is of good quality: such issues as missing values, impossible data combination (e.g., negative number of products), zero values etc. rarely occur. Therefore, the preprocessing step is simpler when dealing with simulation models as compared to real databases.

Manufacturing problems can often be labeled and specialist feedbacks are available, therefore SML techniques are recommended for manufacturing applications (Wuest et al., 2016). The labels in SML may be of discrete or continuous type and can be managed by classification or regression algorithms, respectively (Ribeiro, Grolinger, & Capretz, 2015). The classification is used for prediction, pattern recognition and detection of anomalous values while regression is used for prediction and ranking. Two SML algorithms are used for classification in this work: k-NN and Logistic Regression (LR).

The k-NN algorithm is a non-parametric procedure, i.e., it does not assume prior knowledge of statistical distributions, that assigns to the unclassified instance the nearest instance label using geometric distances (Cover & Hart, 1967). Although LR contains the word regression, it is a learning algorithm used to classify or predict the probability of occurrence of an event by adapting the data to a logistic function and

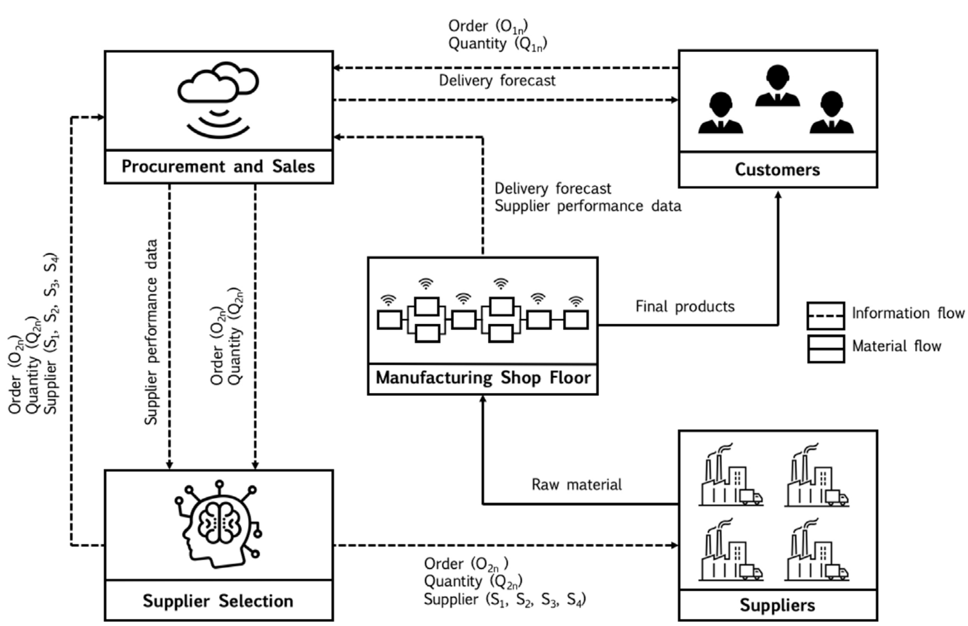


Fig. 1. Make-to-order simulation model.

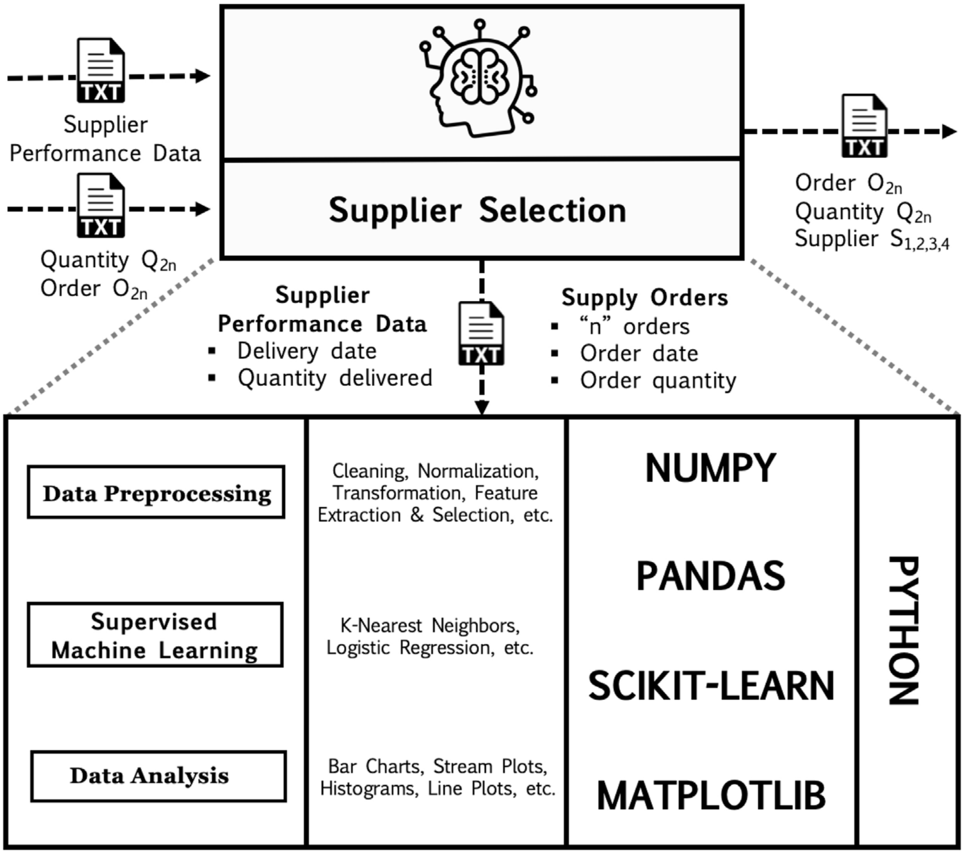


Fig. 2. Supplier selection model using supervised machine learning.

can be used for situations in which the dependent variable is a binary (Yu & Liu, 2011). In addition, LR is a resource that allows estimating the probability associated with the occurrence of a particular event in the face of a set of explanatory variables, i.e., variables that affect the system response and can be defined by the researcher.

The k-NN algorithm is the most common classification algorithm in cases where there is no prior knowledge of data distribution (Li, Kong, Ma, Gong, & Huai, 2016), and LR is based on supervised learning, which organizes itself according to the nature of the input data and there is little need to know about the characteristics of this input data (Yu & Liu, 2011). Although it is known the simulation utilizes normal distribution, in real cases data is likely to differ from a well-behaved normal distribution, so the SML model does not take advantage of any prior knowledge about the system behavior in order to better represent a real case. The aim of the model is to select suppliers with the best chance of delivering an order on time based on past data.

In this work, past data is categorized as (i) deliveries on-time and (ii) late deliveries. The k-NN algorithm is applied separately for each of the two datasets and maps the suppliers' performance according to the previously mentioned characteristics of the model: date and order quantity. In LR, both datasets are the input data and the expected result for each customer order is the probability of each supplier delivering the order within the expected time frame. Therefore, the risk profile represents the probability of success in predicting the supplier behavior in the system regarding the target feature, which is the OTD in this model.

The LR algorithm draws a risk profile for each supplier based on the model's input data, i.e., relevant features that influence the OTD performance for each order: date and order quantity. The output data from this profile is the probability of success in delivering that order on time. After this, through a ranking of the suppliers, the less risky supplier for that particular order is selected. The k-NN algorithm considers the same input data and the algorithm predicts which supplier has the greatest probability of delivering an order on time and which supplier has the greatest probability of performing a late delivery. After this, the supplier with the greatest chance of delivering the order on time is selected.

In addition, a combination of these two techniques is presented in this paper. The first, Hybrid A, confronts the results of both algorithms' classification without considering the accuracy of each technique. The second, Hybrid B, takes the same approach, but considers the accuracy of each classifier.

The accuracy for the k-NN model is the rate in which the model correctly predicts the real outcome. For instance, acc_a and acc_b stand for the accuracy of the k-NN model which uses the deliveries on-time and late deliveries categorizations, respectively. Furthermore, R_{ka} and R_{kb} stand for the k-NN classifier results using the deliveries on-time and late deliveries categorizations, respectively.

The accuracy for the LR model is the area under the receiving operating characteristic (ROC) curve. The area under the curve (AUC) can vary from 0 to 1 and a value of 0.5 is considered a random prediction performance. Fawcett (2006) presents a detailed explanation of ROC curve analysis. Moreover, R_{lr1} and R_{lr2} represent the results for the first and second suppliers most likely to meet the demand on time according to the LR classifier, respectively. Both pseudo-codes are presented as follows.

Algorithm 1. Hybrid A

procedure SELECTION(d, q) ▷ date and quantity
  if $R_{lr1} = R_{kb}$ **then**
return R_{ka}
  else
return R_{lr1}
  end if
end procedure

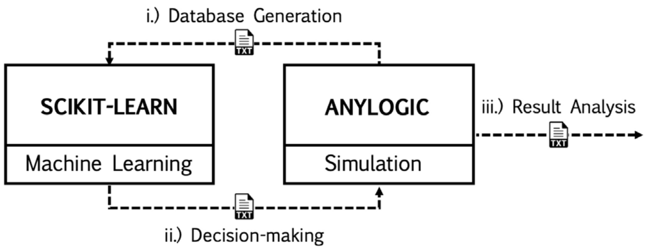


Fig. 3. Integration between simulation and machine learning models.

3.3. Integration

The integration of the simulation and ML models is accomplished through the data exchange results of each model. In this work, the data exchange is achieved with the help of text format files. The sequence of activities for this integration can be summarized in three steps, as shown in Fig. 3.

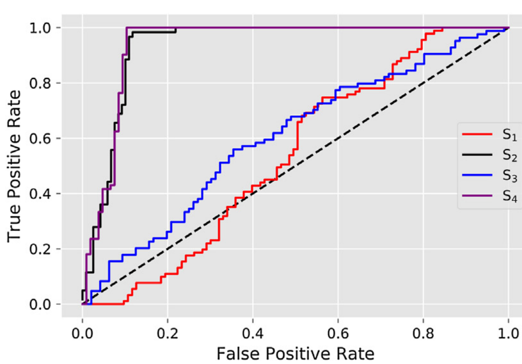
The first step consists of (i) database generation by means of a simulation model. In step two (ii) this database is used as input data in the ML model and then intelligent decision-making results are generated in an output file, which serves as input data to the test simulation experiment. Finally, in step three (iii) the test simulation results are compiled and analyzed.

3.4. Numerical experiment

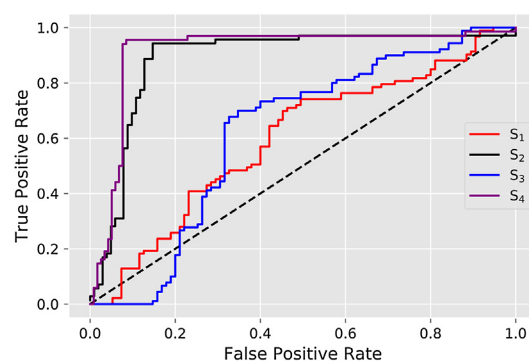
Under the framework shown in Sections 3.1–3.3, we now introduce a scenario under which our modelling methodology would be deployed. In order to evaluate the developed approach, a numerical experiment was conducted. The experiment takes place in a time window of 4 years, in which 50% of the period is used as a training data set, 25% for model validation and tuning and the last 25% as test data. During the training phase, the order allocation to suppliers is random to generate the database of suppliers' performance. Next, SML models train on the training dataset and perform the order allocation to suppliers in test phase. Both algorithms, i.e., LR and k-NN make order allocation towards suppliers which have greatest probability of success in delivering a specific order on time. The performance of the SML models refer only to the results obtained in the test phase.

Algorithm 2. Hybrid B

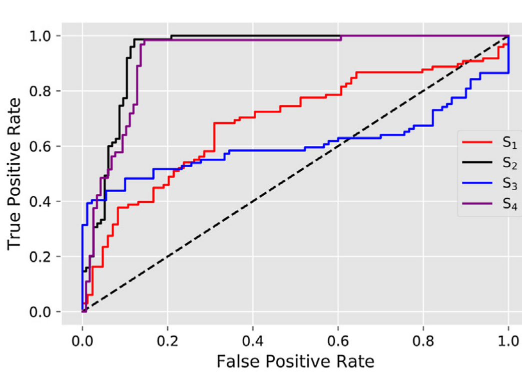
procedure SELECTION(d, q) ▷ date and quantity
  if $AUC \geq acc_a$ **then**
if $AUC \geq acc_b$ **then**
    return R_{lr1}
else
    if $R_{kb} = R_{lr1}$ **then**
return R_{lr2}
    else
return R_{lr1}
    end if
end if
  else
if $AUC \geq acc_b$ **then**
    if $R_{ka} = R_{kb}$ **then**
if $R_{ka} = R_{lr1}$ **then**
    return R_{lr2}
else
    return R_{lr1}
end if
    else
return R_{ka}
end if



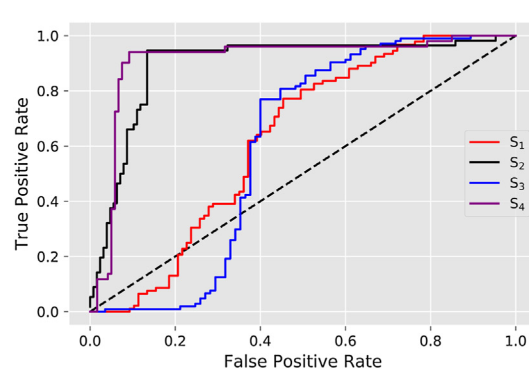
(a) ROC Curve of all suppliers with n_1



(b) ROC Curve of all suppliers with n_2



(c) ROC Curve of all suppliers with n_3



(d) ROC Curve of all suppliers with n_4

Fig. 4. ROC curve of all suppliers (S_1 to S_4).

```
else
return R_ka
end if
  end if
end procedure
```

The first scenario assumes full availability of the four suppliers. The second scenario assumes unavailability of two suppliers due to a disruption in the system, so only two suppliers are available in the testing phase. Therefore, the supplier selection performance will be evaluated by comparing (i) random choice of suppliers, which shows that this kind of data analysis does not exist – correlation between date and order quantity, (ii) using k-NN and (iii) LR algorithms, as well as the combination of these two techniques by means of (iv) Hybrid A and (v) Hybrid B algorithms in both scenarios.

4. Results and discussion

In this section the experiment results are presented using the combination of simulation and SML algorithms for resilient supplier selection. First, the LR performance is shown using four different seeds (n_1 to n_4), as shown in Fig. 4, which translates the prediction accuracy of each supplier given order date and quantity characteristics.

For instance, since the AUC of Supplier 1 and 3 (S_1 and S_3) is inferior to that of the AUC of Supplier 2 and 4 (S_2 and S_4), it is possible to conclude that based on past data, LR predicts the behavior of S_2 and S_4 better than that of S_1 and S_3. Thus, depending on suppliers' characteristics, more accurate models can be found using the same algorithm. In addition, an extract of S_1 and S_2 ROC curves using different simulation seeds is shown in Fig. 5. It can be observed there is a convergence aspect of S_2 compared to S_1. This can be explained by the well defined capacity restriction that has been modeled for S_2 compared to S_1, which makes S_2 more predictable.

Furthermore, a classification sample analyzed via the LR algorithm is shown in Table 1. The algorithm quantifies the probability of each supplier delivering each order on time based exclusively on past data. This approach includes an important aspect of human bias avoidance and the potential to support the decision-making process using other quantitative and qualitative approaches. In this paper, the LR performs the supplier selection both singularly and in combination with k-NN.

As mentioned in Section 3.2, the performance of k-NN algorithm is measured by its accuracy. A set of five different simulation seeds, as presented in Fig. 6, illustrate the accuracy of the k-NN algorithm in this model. In this study, two predictions are made regarding the k-NN model using (i) the on-time delivery categorization and (ii) the delayed delivery categorization. In other words, the model suggests the supplier most likely to deliver a specific order on time and also the supplier most likely to perform a delayed delivery, respectively.

The results show the potential use of SML models as tools for decision-making support. Simulations using different seeds were performed to test the performance of these models based on delivery reliability, which stands proxy for the rate of successful on-time deliveries.

Table 1
LR probability predictions for each supplier and selection results.

Order	S_1	S_2	S_3	S_4	R_{lr1}	R_{lr2}
1	0.4361	0.1972	0.4980	0.1887	S_3	S_1
2	0.3319	0.3402	0.6014	0.3642	S_3	S_4
…	…	…	…	…	…	…
454	0.6236	0.2755	0.3248	0.2475	S_1	S_3
455	0.3313	0.9232	0.6015	0.9493	S_4	S_2
…	…	…	…	…	…	…
729	0.3230	0.0730	0.6108	0.0672	S_3	S_1
730	0.4156	0.2006	0.5178	0.1947	S_3	S_1

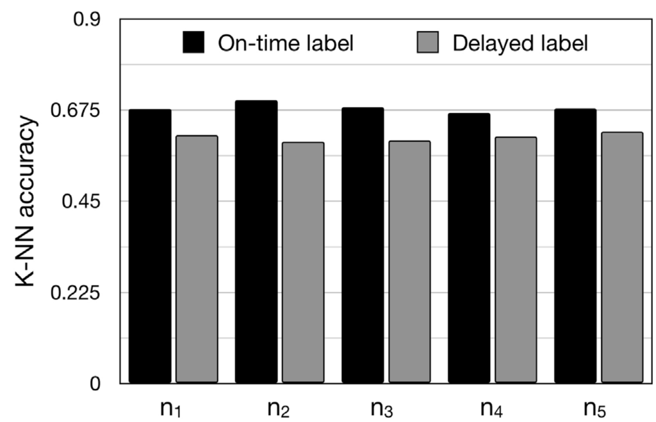
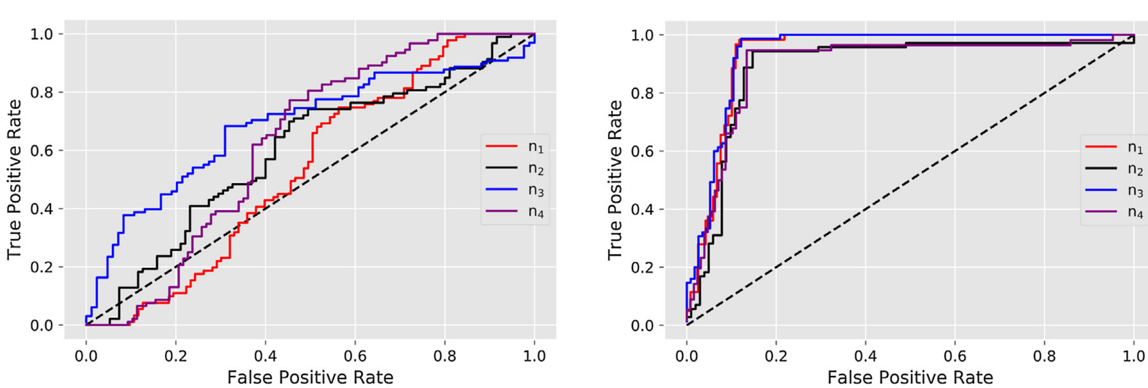


Fig. 6. Accuracy comparison of k-NN model under five different simulation seeds.

Two simulations were performed based on the two previously mentioned scenarios in Section 3.4. Each simulation is repeated using five different seeds and the final results are presented in Fig. 7 according to the mean values.

The experiment results suggest that a higher number of suppliers leads to a more resilient system, which can cope with disruptions and recurrent risks. In part, this is due to the fact there are more assertive models of adequate suppliers for a specific order. More evaluation options to be evaluated are available from which to make good choice. However, it is worth mentioning there is a trade-off – since the number of orders in the period does not change, a higher number of suppliers leads to less data being analyzed for each of them. For instance, in this experiment the total number of orders is 2921. Thus, with less known



(a) ROC Curve of S_1 (b) ROC Curve of S_2

Fig. 5. ROC curve of S_1 and S_2 after simulation with four different seeds (n_1 to n_4).

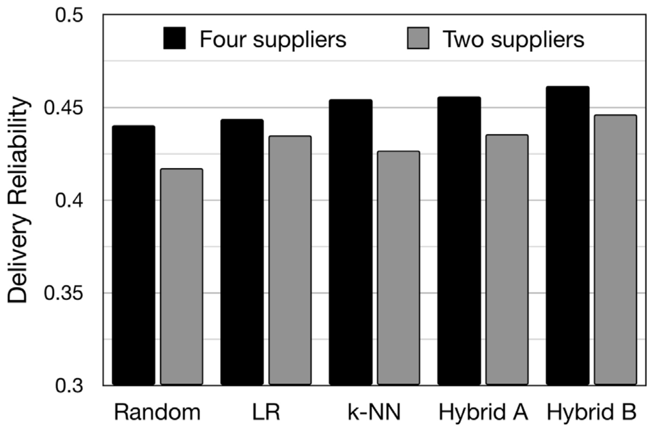


Fig. 7. Delivery reliability performance after supplier selection using supervised machine learning.

information about a given supplier, there is a tendency towards poor representation of reality by the model, which means less accuracy, followed by poorer results with which to obtain a resilient supplier selection.

It can also be observed that the mixed use of the SML algorithms led to an improvement in the delivery reliability of suppliers. For example, using the random approach to supplier selection, meaning this kind of data is not analyzed by SML, the delivery reliability is 44.03%. Adopting the Hybrid B model, the result increases to 46.16%. This means 62 late deliveries were avoided with the simple use of data that a priori would not be analyzed.

When generalizing the results of this study, it can be observed that Big Data is worthless if not leveraged to drive decision-making (Gandomi & Haider, 2015). In a society increasingly influenced by data-driven decision-making, the use of any and all kinds of data have the potential to generate new forms of decision-making and negotiation mechanisms. Emerging services and analytics, including merged technologies such as data warehouse, ML, visualization (Larson & Chang, 2016), are a new form of value creation in the era of digital manufacturing. Therefore, a complex and disruptive reality emerges and strategic and tactical decisions must consider the impact of digital fragmentation in all aspects of the business (Abbosh, Nunes, & Ovanessoff, 2017), including the fragmentation of relationships between manufacturing firms.

The experiment results show that a combination of SML and simulation can help in specifying a risk profile for each order, i.e., based on two features (delivery time and quantity) that have a causality relationship to OTD. The SC managers can obtain the estimations of what suppliers, or combinations of suppliers are most critical in terms of the disruptions and the resulting SC performance impact. As such the managers can explicitly use a causality relationship of the parameters in risk profiles of supplier performance with OTD (or any other KPI) that in turn, could feed risk mitigation and building resilient supplier profiles. These risk profiles, which are built based on past data, can help creating continuous improvement strategies for supplier portfolio development.

Ad-hoc customer-supplier relationships may arise from the adoption of a data-driven culture in manufacturing firms. For instance, a data-driven culture affects the bargaining power of companies, which could be represented by smart contracts based on supplier selection predictive models. In addition, by collecting and analyzing performance data from suppliers, it could be possible to contribute to more robust risk management models, which in turn would increase SC resilience.

Digital manufacturing points to the direction of convergence between real and digital worlds by means of massive use of data and digital twins drive agile experimentation to enhance production systems. Experience-based decision-making tends to be replaced or, at least, strongly supported by data-driven decision-making. A priori, the resilient supplier selection has a role of decision-making support since it is an intrinsic multi-criteria decision problem and must consider strategic decisions. Therefore, the adoption of resilient supplier selection in combination with the strategic decision-making level has potential to compose a robust system of supplier selection in a digital manufacturing environment.

5. Managerial implications and theoretical contributions

Several managerial implications can be highlighted from this work, these insights may pose directions to future practical implementations to resilient SC development. First, our analysis shows an information method to integrate simulation and ML models that can evaluate digital services performance in manufacturing. Since it is a fully digital approach, it can be valuable for prototyping and validating new services in less time and cost within digital manufacturing context. Second, the adoption of a data-driven culture in manufacturing enterprises may result in ad-hoc customer-supplier relationships. This can happen due to the possibility of developing bias-free ML models, which means decision-making can be exclusively result-oriented. Third, the model utilizes data that does not require any expensive data acquisition system, therefore this kind of approach can be seen essential to increment the rate of early adopters of digital manufacturing. In addition, data management must consider strategic decisions to unlock benefits and develop data strategies within manufacturing firms. Forth, intelligent and agile decision-making is considered essential to develop resilient SCs, therefore digital twins are useful to prospect scenarios in order to achieve resilient systems by performing proactive agile experimentation.

Furthermore, some theoretical contributions are emphasized. First, the use of SML based on existing databases may boost SC risk management models. This can happen because the use of SML models allow the reduction of abstractions of risk management models by analysing past data and presenting pattern recognition outcomes that can substitute diverse simplifying hypothesis. Second, rule-based systems combining learning algorithms can increment overall system performance. In this work, two algorithms (Hybrid A and B) improved overall delivery reliability by manipulating the accuracy of learning algorithms. Third, as the proposed model is based on a learning process, it has potential to confer adaptability to the decision-making process and can dynamically analyse past data in order to make better decisions. Forth, previous researches using ML to solve supplier selection do not have presented simulation approaches, which are likely to gain momentum with the digitization of manufacturing assets by IoT devices. Therefore, this work contributes to the vision of using manufacturing simulation in a new way, i.e., as a provider of synthetic data to train ML models that address SC resilience. Finally, manufacturing is becoming increasingly dependent on statistical methods and there is a wide variety of data analytics approaches that could be experimented to solve classical manufacturing problems.

6. Conclusion

In this paper, we introduced a new approach to resilient supplier selection that utilizes the advances in data analytics while avoiding two major inconveniences, namely the need to estimate the likelihood of disruptions and forecasting the performance impacts. One difficulty in managing the resilient supplier portfolios using disruption probability estimations is a relative rarity of risk events which are too intermittent and irregular to be accurately identified, estimated, and forecasted. Instead of estimating probabilities of highly unpredictable events, the emphasis of our study shifts to utilizing the advantages of digital data in smart manufacturing systems to predict the supplier proneness to disruptions, and the associated impact on SC performance. A specific focus of analysis was directed toward resilient supplier selection in digital manufacturing. The test cases were performed in a digital make-to-order manufacturing environment using a simulation tool. The results indicate that the use of SML algorithms can support the resilient supplier selection decision-making process, leading to more predictable delivery from suppliers and improvements in risk mitigation decision-making. The application of this approach requires a change of mindset regarding the customer-supplier relationship, meaning that these relations should be more ephemeral and data-oriented so that resilient supplier portfolios can be developed and resilient SCs can be achieved.

Two significant contributions emerge. First, we show that the associations of the deviations from the resilient SC performance profile with the risk profiles of supplier performance can be efficiently deciphered by a combination of SML and simulation. Second, the results of this study advance the understanding about how and when ML and simulation can be combined to create digital SC twins, and through these twins improve resilience. The outcomes of this study can emerge in a number of useful insights for managers such as a development of most critical suppliers, re-engineering of supplier base, investments in SC resilience, order allocation improvement or even an acquisition of a risky but very important supplier. The findings suggest that our model can be of value in revealing latent, high-risk supplier portfolios, and prioritizing risk mitigation efforts. In the experiment, the suppliers had restrictions on production capacity in certain periods and were represented in a dataset divided by categories, such as order date and order quantity. The SML model was able to predict the performance of the suppliers when variations in these categorizations had occurred.

The use of SML can contribute to supplier selection as a risk mitigation strategy that could assist optimization and resilience management models. With the advent of Big Data availability, decision-making in manufacturing will become increasingly dependent on statistical methods. Hence, it is essential to pave the way for replacing abstractions with ML models in manufacturing risk management processes, so that value creation can be perceived by practitioners and real data shared, leading to a virtuous circle of improvement.

Finally, some limitations and future research avenues may be highlighted. First, the advantages of using ML techniques can become more evident when considering larger data sets. Those advantages can be manifested in faster processing times and better causality recognition as compared to traditional statistical methods. Since the dimensionality of our data set is quite small and restricted to two parameters (i.e., delivery time and quantity), other statistical methods could have been used for our specific model, but on the other hand, such methods could not be feasible in real applications. In real supplier databases, there would be multiple parameters in the SC resilience analysis. The use of ML could suit better to such an increased complexity and can be of value at manufacturing firms with a data-driven culture. Second, although the model considers stochastic variations to approximate to a real case, the model is still based on fictional data: the results are subject to variations in real case scenarios. For real case applications in data-oriented firms, more features will exist because of the increase in data availability. In these cases, previous feature selection can be used to identify the most relevant features in the prediction model, or deep learning techniques should be considered. To that end, the simulation model can be extended by adding product variability, transport costs, and other customized features. In addition, it is possible to investigate different SML algorithms, as well as new methods of combining two or more of these algorithms while considering the respective accuracy of each.

These limitations imply a number of possible extensions of this work in future. For example, a differentiation of supplier profiles can be considered, e.g., a more resilient supplier has higher costs, or a variation in available quantity is different at different suppliers, or a price competition between suppliers. Furthermore, the use of rule-based systems combining different learning algorithms showed overall system performance improvement. This may be an indicative that the use of learning subsystems via meta-learning may yield even better performances specially when modelling in more complex scenarios. These extensions would also be favorable by introducing other methodological aspects, e.g., deep learning techniques which might be helpful in detecting multiple causalities and improving the model performance.

References

Abbosh, O., Nunes, P., & Ovanessoff, A. (2017). Adapting your digital business to a fragmented world. *Harvard Business Review, 95*, 56–61.
Akter, S., Bandara, R., Hani, U., Wamba, S. F., Foropon, C., & Papadopoulos, T. (2019). Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management, 48*, 85–95.
Altay, N., Gunasekaran, A., Dubey, R., & Childe, S. J. (2018). Agility and resilience as antecedents of supply chain performance under moderating effects of organizational culture within the humanitarian setting: A dynamic capability view. *Production Planning & Control, 29*, 1158–1174.
Ang, E., Iancu, D. A., & Swinney, R. (2016). Disruption risk and optimal sourcing in multitier supply networks. *Management Science, 63*, 2397–2419.
Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2018). Supply chain risk management and artificial intelligence: State of the art and future research directions. *International Journal of Production Research,* 1–24.
Bodaghi, G., Jolai, F., & Rabbani, M. (2018). An integrated weighted fuzzy multi-objective model for supplier selection and order scheduling in a supply chain. *International Journal of Production Research, 56*, 3590–3614.
Bohner, C., & Minner, S. (2017). Supplier selection under failure risk, quantity and business volume discounts. *Computers & Industrial Engineering, 104*, 145–155.
Bounfour, A. (2016). *Digital futures, digital transformation.* Springer.
Carrasco, U., Coronado, P. D. U., Parto, M., & Kurfess, T. (2018). Indoor location service in support of a smart manufacturing facility. *Computers in Industry, 103*, 132–140.
Chae, B. K. (2019). A general framework for studying the evolution of the digital innovation ecosystem: The case of big data. *International Journal of Information Management, 45*, 83–94.
Chen, A., Hsieh, C.-Y., & Wee, H. (2016). A resilient global supplier selection strategy – A case study of an automotive company. *The International Journal of Advanced Manufacturing Technology, 87*, 1475–1490.
Chen, Y.-J. (2011). Structured methodology for supplier selection and evaluation in a supply chain. *Information Sciences, 181*, 1651–1670.
Choi, T.-M. (2018). A system of systems approach for global supply chain management in the big data era. *IEEE Engineering Management Review, 46*, 91–97.
Choi, T.-M., Chan, H. K., & Yue, X. (2017). Recent development in big data analytics for business operations and risk management. *IEEE Transactions on Cybernetics, 47*, 81–92.
Choi, T.-M., & Lambert, J. H. (2017). Advances in risk analysis with big data. *Risk Analysis, 37*, 1435–1442.
Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management, 27*, 1868–1883.
Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*, 21–27.
Das, A., Narasimhan, R., & Talluri, S. (2006). Supplier integration – Finding an optimal configuration. *Journal of Operations Management, 24*, 563–582.
Dey, P. K., Bhattacharya, A., Ho, W., & Clegg, B. (2015). Strategic supplier performance evaluation: A case-based action research of a UK manufacturing organisation. *International Journal of Production Economics, 166*, 192–214.
Diaz-Rozo, J., Bielza, C., & Larra naga, P. (2017). Machine learning-based CPS for clustering high throughput machining cycle conditions. *Procedia Manufacturing, 10*, 997–1008.
Dickson, G. W. (1966). An analysis of vendor selection systems and decisions. *Journal of Purchasing, 2*, 5–17.
Dolgui, A., Bakhtadze, N., Pyatetsky, V., Sabitov, R., Smirnova, G., Elpashev, D., et al. (2018a). Data mining-based prediction of manufacturing situations. *IFAC-PapersOnLine, 51*, 316–321.
Dolgui, A., Ivanov, D., & Sokolov, B. (2018b). Ripple effect in the supply chain: An analysis and recent literature. *International Journal of Production Research, 56*, 414–430.
Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data-evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.
Dubey, R., Gunasekaran, A., & Childe, S. J. (2015). The design of a responsive sustainable supply chain network under uncertainty. *The International Journal of Advanced Manufacturing Technology, 80*, 427–445.
Dubey, R., Gunasekaran, A., Childe, S., Fosso Wamba, S., Roubaud, D., & Forupon, C. (2019). Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *International Journal of Production Research*.
Dupont, L., Bernard, C., Hamdi, F., & Masmoudi, F. (2018). Supplier selection under risk of delivery failure: A decision-support model considering managers risk sensitivity. *International Journal of Production Research, 56*, 1054–1069.
Durugbo, C., Tiwari, A., & Alcock, J. R. (2013). Modelling information flow for organisations: A review of approaches and future challenges. *International Journal of Information Management, 33*, 597–610.
Elluru, S., Gupta, H., Kaur, H., & Singh, S. P. (2017). Proactive and reactive models for disaster resilient supply chain. *Annals of Operations Research,* 1–26.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*, 861–874.
Frazzon, E. M., Kück, M., & Freitag, M. (2018). Data-driven production control for complex and dynamic manufacturing systems. *CIRP Annals*.
Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*, 137–144.
Gao, S. Y., Simchi-Levi, D., Teo, C.-P., & Yan, Z. (2018). Disruption risk mitigation in supply chains – The risk exposure index revisited. *Operations Research*.
Gosling, J., Purvis, L., & Naim, M. M. (2010). Supply chain flexibility as a determinant of supplier selection. *International Journal of Production Economics, 128*, 11–21.
Grant, D., & Yeo, B. (2018). A global perspective on tech investment, financing, and ICT on manufacturing and service industry performance. *International Journal of Information Management, 43*, 130–145.
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*, 525–527.
Guo, X., Yuan, Z., & Tian, B. (2009). Supplier selection based on hierarchical potential support vector machine. *Expert Systems with Applications, 36*, 6978–6985.
Gusikhin, O., & Klampfl, E. (2012). *JEDI: Just-in-time execution and distribution information support system for automotive stamping operations. Decision policies for production networks*. Springer119–142.
Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., & Monostori, L. (2018). Lead time prediction in a flow-shop environment with analytical and machine learning approaches. *IFAC-PapersOnLine, 51*, 1029–1034.
Hamdi, F., Ghorbel, A., Masmoudi, F., & Dupont, L. (2018). Optimization of a supply portfolio in the context of supply chain risk management: Literature review. *Journal of Intelligent Manufacturing, 29*, 763–788.
He, J., Alavifard, F., Ivanov, D., & Jahani, H. (2018). A real-option approach to mitigate disruption risk in the supply chain. *Omega, 08*, 008 https://doi.org/10.1016/j.
Hosseini, S., & Barker, K. (2016). A Bayesian network model for resilience-based supplier selection. *International Journal of Production Economics, 180*, 68–87.
Hosseini, S., Ivanov, D., & Dolgui, A. (2019a). Review of quantitative methods for supply chain resilience analysis. *Transportation Review: Part E, 125*, 285–307.
Hosseini, S., Morshedlou, N., Ivanov, D., Sarder, M. D., Barker, K., & Al Khaled, A. (2019b). Resilient sup-plier selection and optimal order allocation under disruption risks. *International Journal of Production Economics, 03*, 018. https://doi.org/10.1016/j.ijpe.2019.03.018.
Ismagilova, E., Hughes, L., Dwivedi, Y. K., & Raman, K. R. (2019). Smart cities: Advances in research—An information systems perspective. *International Journal of Information Management, 47*, 88–100.
Ivanov, D. (2018). *Structural dynamics and resilience in supply chain risk management.* Springer.
Ivanov, D. (2019). *Managing risks in supply chains with digital twins and simulation, White paper*.
Ivanov, D., & Dolgui, A. (2018). Low-Certainty-Need (LCN) supply chains: A new perspective in managing disruption risks and resilience. *International Journal of Production Research,* 1–18.
Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research, 57*(3), 829–846.
Ivanov, D., Dolgui, A., Sokolov, B., & Ivanova, M. (2017). Literature review on disruption recovery in the supply chain. *International Journal of Production Research, 55*, 6158–6174.
Ivanov, D., Tsipoulanidis, A., & Schönberger, J. (2017). *Global supply chain and operations management. A decision-oriented introduction to the creation of value*.
Ji, W., & Wang, L. (2017). Big data analytics based fault prediction for shop floor scheduling. *Journal of Manufacturing Systems, 43*, 187–194.
Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*, 255–260.
Jüttner, U., Peck, H., & Christopher, M. (2003). Supply chain risk management: Outlining an agenda for future research. *International Journal of Logistics: Research and Applications, 6*, 197–210.
Karim, M. A., Samaranayake, P., Smith, A., & Halgamuge, S. K. (2010). An on-time delivery improvement model for manufacturing organisations. *International Journal of Production Research, 48*, 2373–2394.
Kartal, H., Oztekin, A., Gunasekaran, A., & Cebi, F. (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. *Computers & Industrial Engineering, 101*, 599–613.
Kaur, H., & Singh, S. P. (2016). Sustainable procurement and logistics for disaster resilient supply chain. *Annals of Operations Research,* 1–46.
Kaur, H., & Singh, S. P. (2017). Modeling low carbon procurement and logistics in supply chain: A key towards sustainable production. *Sustainable Production and Consumption, 11*, 5–17.
Kaur, H., & Singh, S. P. (2018). Heuristic modeling for sustainable procurement and logistics in a supply chain using big data. *Computers & Operations Research, 98*, 301–321.
Kho, D. D., Lee, S., & Zhong, R. Y. (2018). Big data analytics for processing time analysis in an IoT-enabled manufacturing shop floor. *Procedia Manufacturing, 26*, 1411–1420.
Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Data preprocessing for supervised leaning. *International Journal of Computer Science, 1*, 111–117.
Kruger, G. H., Shih, A. J., Hattingh, D. G., & van Niekerk, T. I. (2011). Intelligent machine agent architecture for adaptive control optimization of manufacturing processes. *Advanced Engineering Informatics, 25*, 783–796.
Kshetri, N. (2018). Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management, 39*, 80–89.
Kull, T. J., & Talluri, S. (2008). A supply risk reduction model using integrated multicriteria decision making. *IEEE Transactions on Engineering Management, 55*, 409–419.
Kumar, A., Mangla, S. K., Luthra, S., Rana, N. P., & Dwivedi, Y. K. (2018). Predicting changing pattern: Building model for consumer decision making in digital market. *Journal of Enterprise Information Management, 31*, 674–703.
Kumar, R., Singh, S. P., & Lamba, K. (2018). Sustainable robust layout using big data approach: A key towards Industry 4.0. *Journal of Cleaner Production, 204*, 643–659.
Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature News, 544*, 23.
Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research, 56*, 508–517.
Lamba, K., & Singh, S. P. (2017). Big data in operations and supply chain management: Current trends and future perspectives. *Production Planning & Control, 28*, 877–890.
Lamba, K., & Singh, S. P. (2018). Dynamic supplier selection and lot-sizing problem considering carbon emissions in a big data environment. *Technological Forecasting and Social Change*.
Lamba, K., Singh, S. P., & Mishra, N. (2019). Integrated decisions for supplier selection and lot-sizing considering different carbon emission regulations in big data environment. *Computers & Industrial Engineering, 128*, 1052–1062.
Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management, 36*, 700–710.
Li, N., Kong, H., Ma, Y., Gong, G., & Huai, W. (2016). Human performance modeling for manufacturing based on an improved KNN algorithm. *The International Journal of Advanced Manufacturing Technology, 84*, 473–483.
Li, X., Wang, J., & Sawhney, R. (2012). Reinforcement learning for joint pricing, lead-time and scheduling decisions in make-to-order systems. *European Journal of Operational Research, 221*, 99–109.
Liao, Y., Deschamps, F., Loures, E.d. F. R., & Ramos, L. F. P. (2017). Past, present and future of Industry 4.0—A systematic literature review and research agenda proposal. *International Journal of Production Research, 55*, 3609–3629.
Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., & Monostori, L. (2018). Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. *Procedia CIRP, 72*, 1051–1056.
Liu, S., Chan, F. T., Yang, J., & Niu, B. (2018). Understanding the effect of cloud computing on organizational agility: An empirical examination. *International Journal of Information Management, 43*, 98–111.
Mirkouei, A., & Haapala, K. R. (2014). Integration of machine learning and mathematical programming methods into the biomass feedstock supplier selection process. *Flexible Automation and Intelligent Manufacturing*.
Narasimhan, R., & Talluri, S. (2009). Perspectives on risk management in supply chains. *Journal of Operations Management, 27*, 114–118.
Nguyen, T., Li, Z., Spiegler, V., Ieromonachou, P., & Lin, Y. (2018). Big data analytics in supply chain management: A state-of-the-art literature review. *Computers & Operations Research, 98*, 254–264.
Palombarini, J., & Martínez, E. (2012). SmartGantt – An intelligent system for real time rescheduling based on relational reinforcement learning. *Expert Systems with Applications, 39*, 10251–10268.
Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S. J., & Fosso-Wamba, S. (2017). The role of Big Data in explaining disaster resilience in supply chains for sustainability. *Journal of Cleaner Production, 142*, 1108–1118.
Pavlov, A., Ivanov, D., Pavlov, D., & Slinko, A. (2019). Optimization of network redundancy and contingency planning in sustainable and resilient supply chain resource management under conditions of structural dynamics. *Annals of Operations Research*.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
PrasannaVenkatesan, S., & Goh, M. (2016). Multi-objective supplier selection and order allocation under disruption risk. *Transportation Research Part E: Logistics and Transportation Review, 95*, 124–142.
Priore, P., Ponte, B., Puente, J., & Gómez, A. (2018). Learning-based scheduling of flexible manufacturing systems using ensemble methods. *Computers & Industrial Engineering, 126*, 282–291.
Rajagopal, V., Venkatesan, S. P., & Goh, M. (2017). Decision-making models for supply chain risk mitigation: A review. *Computers & Industrial Engineering, 113*, 646–682.
Rajesh, R., & Ravi, V. (2015). Supplier selection in resilient supply chains: A grey relational analysis approach. *Journal of Cleaner Production, 86*, 343–359.
Rana, N. P., Luthra, S., Mangla, S. K., Islam, R., Roderick, S., & Dwivedi, Y. K. (2018). Barriers to the development of smart cities in Indian context. *Information Systems Frontiers,* 1–23.
Ribeiro, M., Grolinger, K., & Capretz, M. A. (2015). MLaaS: Machine learning as a service. *2015 IEEE 14th international conference on machine learning and applications (ICMLA)* (pp. 896–902).
Rossit, D. A., Tohmé, F., & Frutos, M. (2018). Industry 4.0: Smart scheduling. *International Journal of Production Research,* 1–12.
Sawik, T. (2013a). Selection of resilient supply portfolio under disruption risks. *Omega, 41*, 259–269.
Sawik, T. (2013b). Integrated selection of suppliers and scheduling of customer orders in the presence of supply chain disruption risks. *International Journal of Production Research, 51*, 7006–7022.
Sawik, T. (2016). On the risk-averse optimization of service level in a supply chain under disruption risks. *International Journal of Production Research, 54*, 98–113.
Scheibe, K. P., & Blackhurst, J. (2018). Supply chain disruption propagation: A systemic risk and normal accident theory perspective. *International Journal of Production Research, 56*, 43–59.
Shahzad, A., & Mebarki, N. (2012). Data mining based job dispatching using hybrid

simulation-optimization approach for shop scheduling problem. *Engineering Applications of Artificial Intelligence, 25*, 1173–1181.
Sharp, M., Ak, R., & Hedberg, T., Jr. (2018). A survey of the advancing use and development of machine learning in smart manufacturing. *Journal of Manufacturing Systems*.
Sheffi, Y., & Rice, J. B., Jr. (2005). A supply chain view of the resilient enterprise. *MIT Sloan Management Review, 47*, 41.
Simchi-Levi, D., Schmidt, W., Wei, Y., Zhang, P. Y., Combs, K., Ge, Y., et al. (2015). Identifying risks and mitigating disruptions in the automotive supply chain. *Interfaces, 45*, 375–390.
Solti, A., Raffel, M., Romagnoli, G., & Mendling, J. (2018). Misplaced product detection using sensor data without planograms. *Decision Support Systems, 112*, 76–87.
Stricker, N., Kuhnle, A., Sturm, R., & Friess, S. (2018). Reinforcement learning for adaptive order dispatching in the semiconductor industry. *CIRP Annals*.
Strozzi, F., Colicchia, C., Creazza, A., & Noè, C. (2017). Literature review on the 'Smart Factory' concept using bibliometric tools. *International Journal of Production Research, 55*, 6572–6591.
Tomlin, B. (2006). On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science, 52*, 639–657.
Torabi, S., Baghersad, M., & Mansouri, S. (2015). Resilient supplier selection and order allocation under operational and disruption risks. *Transportation Research Part E: Logistics and Transportation Review, 79*, 22–48.
Tuncel, E., Zeid, A., & Kamarthi, S. (2014). Solving large scale disassembly line balancing problem with uncertainty using reinforcement learning. *Journal of Intelligent Manufacturing, 25*, 647–659.
Viswanadham, N., & Samvedi, A. (2013). Supplier selection based on supply chain ecosystem, performance and risk criteria. *International Journal of Production Research, 51*, 6484–6498.
Vugrin, E. D., Warren, D. E., & Ehlen, M. A. (2011). A resilience assessment framework for infrastructure and economic systems: Quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane. *Process Safety Progress, 30*, 280–290.
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356–365.
Wang, J., Yang, J., Zhang, J., Wang, X., & Zhang, W. (2018a). Big data driven cycle time parallel prediction for production planning in wafer manufacturing. *Enterprise Information Systems, 12*, 714–732.
Wang, J., Zhang, J., & Wang, X. (2018b). Bilateral LSTM: A two-dimensional long short-term memory model with multiply memory units for short-term cycle time forecasting in re-entrant manufacturing systems. *IEEE Transactions on Industrial Informatics, 14*, 748–758.
Waschneck, B., Reichstaller, A., Belzner, L., Altenmüller, T., Bauernhansl, T., Knapp, A., et al. (2018). Optimization of global production scheduling with deep reinforcement learning. *Procedia CIRP, 72*, 1264–1269.
Weber, C. A., & Current, J. R. (1993). A multiobjective approach to vendor selection. *European Journal of Operational Research, 68*, 173–184.
Wetzstein, A., Hartmann, E., Benton, W., Jr., & Hohenstein, N.-O. (2016). A systematic assessment of supplier selection literature—State-of-the-art and future scope. *International Journal of Production Economics, 182*, 304–323.
Wieland, A., & Marcus Wallenburg, C. (2013). The influence of relational competencies on supply chain resilience: A relational view. *International Journal of Physical Distribution & Logistics Management, 43*, 300–320.
Wu, D., & Olson, D. L. (2008). Supply chain risk, simulation, and vendor selection. *International Journal of Production Economics, 114*, 646–655.
Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research, 4*, 23–45.
Yoon, J., Talluri, S., Yildiz, H., & Ho, W. (2018). Models for supplier selection and risk mitigation: A holistic approach. *International Journal of Production Research, 56*, 3636–3661.
Yu, J., & Liu, J. (2011). LRProb control chart based on logistic regression for monitoring mean shifts of auto-correlated manufacturing processes. *International Journal of Production Research, 49*, 2301–2326.