Data Science supporting Lean production: a bibliometric study

Cannas V. G.*, Ciano M. P.*, Pozzi R.*, Rossi T.*

** Scuola di Ingegneria Industriale, Univerisità Carlo Cattaneo LIUC, Corso Matteotti 22, 21053 – Castellanza – Italy (vcannas@liuc.it, mciano@liuc.it, rpozzi@liuc.it, trossi@liuc.it)*

Abstract: In the Industry 4.0 era, the amount of available data is rapidly increasing, and data science shows promise in the manufacturing transformation. Data science can potentially boost lean production tools, facilitating their implementation and enhancing their effects through the use of statistics, mathematics, computer and information science to extract useful knowledge. Notwithstanding the known and potential advantages of its adoption, data science is hardly exploited by manufacturers as they often do not know what the right tool is to achieve the desired objective. As this topic is still under-investigated, this paper aims at filling the gap proposing a review of literature to support the exploitation of data science in lean production contexts. A quantitative bibliometric method is applied to define what tools and techniques have been adopted within the context of lean production. This study demonstrates that the topic has been experiencing a real development only in recent years, and data science elements have been adopted by literature mainly to make inferences about lean. Additionally, clustering techniques were applied to both indexed and author co-occurred keywords and 11 clusters were identified. From the analysis of clusters, this work identifies the principles and industries that have been mainly interested by the use of data, providing a knowledge base on developed topics and underrepresented ones. The obtained map of previous research would act as a guide for managers that seek to boost the positive effects of lean production.

Keywords: Lean production, data science, bibliometric analysis, co-occurrence network

1.Introduction

Lean production aims at reducing waste, improving the efficiency, i.e., costs and lead time reduction, and increasing the effectiveness, i.e., quality enhancement, of the manufacturing processes (Womack et al., 2007). Its implementation has recently been pointed as a prerequisite for the adoption of digital technologies (Buer et al., 2018). In the Industry 4.0 era, the amount of available data is rapidly increasing; massive data are generated continuously with the production activities and contain very useful information (Zhong et al., 2017). Data science has experienced several developments in the last fifty years, but according to Donoho (2017) the largest vision of data science can be classified into six activities: (i) data gathering, preparation, and exploration; (ii) data representation and transformation; (iii) computing with data; (iv) data modeling; (v) data visualization and presentation; (vi) and finally science itself. Therefore, data science shows promise in the manufacturing transformation, since it allows to extract information from data (Donoho, 2017). Data-driven decision making helps companies in getting competitive advantage and has high potential for achieving excellence (Rejikumar et al., 2020). For this reason, data science has been pointed as a mean to potentially boost lean production tools (Buer et al., 2018). For instance, data mining techniques could help in identifying patterns or other machine learning models to map predictors in datasets related to machines behaviours and enhance total productive maintenance activities.

Nowadays, the interest on the analytical value of data for analysing, predicting and controlling processes in business and industry, is increasing. However, despite the known and potential advantages of its adoption, data science is hardly exploited by manufacturers as they often do not know what the right tool is to achieve the desired objective (Coleman et al., 2016). As this topic is still under-investigated, this paper aims at filling the gap by answering to the following research question:

RQ. *What data science activities have been used in the lean production context?*

To achieve this aim, the study proposes a review of literature (materials are presented in Sec. 2) through a quantitative bibliometric method (outlined in Sec. 3), based on the keywords co-occurrence network, applied to a set of papers extracted from Scopus. The study (outlined in Sec. 4 and 5 and discussed in Sec. 6) contribute to literature as a first attempt to map the most relevant papers about the topic and identifythe lean areas that have already experienced implementations and that seem to be promising and propose future directions (as underlined by Sec. 7).

2.Materials

The dataset for this study consists of papers collected in Scopus, the most extensive database of scientific peer-reviewed production (Pozzi and Strozzi 2018). The search was performed using the TITLE-ABS-KEY field, where ABS is a contract form for abstract, and the KEY field includes AUTHKEY (author keywords) and different kinds of indexed keywords. Author keywords are keywords chosen by the authors themselves to describe the specific content of their work. Indexed keywords are vocabulary and thesaurus terms that content suppliers addressed to a publication in order to embrace all of its characteristics in a broader and more comprehensive way (Pozzi and Strozzi 2018; Ciano et al. 2019). The use of the ABS field can include papers not precisely devoted to the link between lean and data science, but whose abstract contain a reference to specific parts of them in which the two topics were combined. Therefore, the use of all these fields can encompass more works.

Lean and data science can have several subtopics, tools, practices, and synonyms.
Referring to data science, all terms introduced by Donoho (2017) were considered in the search. In addition, "big data" and "data anal*" (with the possible ends in -ysis or -ytics) were added. Indeed, "big data" and data science are strictly linked since the latter is the necessary tool for handling the formers (Donoho 2017). The second term was included because data science can be seen as the field that studies "data analysis" (Cleveland 2001; Donoho 2017) and "data analytics" as the set of models and algorithms for intelligent data analysis (Runkler 2012).

Regarding lean, the article by Ciano et al. (2019) suggested that the literature about the topic can be encompassed searching for the terms "lean", "toyota production system", "just in time", "six sigma", "total quality management" and all their acronyms or contracted forms. Therefore, the following string is the search performed on Scopus: (TITLE-ABS-KEY ("data science" OR "big data" OR "data Gathering" OR "data Preparation" OR "data Exploration" OR "data Representation" OR "data Transformation" OR "data Computing" OR "data Modeling" OR "data Visualization" OR "data Presentation" OR "data anal*") AND (TITLE-ABS-KEY ("lean" OR "tps" OR "jit" OR "just in time" OR "just-in-time" OR "tqm" OR "six sigma" OR "six-sigma" OR "toyota production system" OR "total quality management"))). Another final screening was adopted excluding all the subject areas not related to engineering, management, economics, computer, and decision science.

In February 2019, the output of this search was a set of 593 papers distributed over time as depicted in Figure 1. The figure reveals that the scientific production about the topic has been experiencing a general growing trend since the early 2000s and a sudden increase in the last five years.

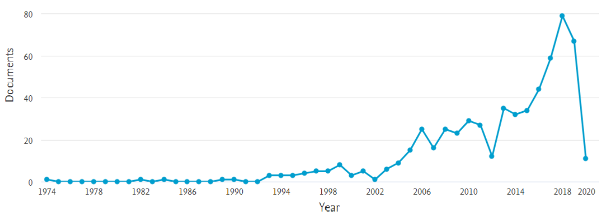

Figure 1: documents by year

3.Method

This paper tackles the research question through a bibliometric analysis. In particular, the method consists of two elements of the Systematic Literature Network Analysis (SLNA) (Colicchia and Strozzi 2012; Strozzi et al. 2017; Ciano et al. 2019), namely the publications' main path analysis and the keywords' co-occurrence network analysis (Table 1).

The publications' main path (figure 2) is an extraction of the publications' citation network. The extraction is possible considering only connected components. In the citation network, the connected components, namely nodes representing publications linked by citations, depend on prior works that have influenced their content. Their links are graphically depicted by arrows from the citing paper to the cited one, representing the flow of knowledge. The main path represents the main development trajectory of the field, the "backbone of the research tradition" (Lucio-Arias and Leydesdorff 2008, Colicchia and Strozzi 2012), and it "highlights the articles that build on prior articles but continue to act as hubs in reference to later works" (Strozzi et al. 2017, 5). Specifically, the approach followed in this research is the detection of the key-route main path, where the key-route is the link that has the highest significance, measured through the so-called "traversal count" (Liu and Lu 2012). This approach was implemented in Pajek, a software for the visualisation and the analysis of large networks (http://mrvar.fdv.uni-lj.si/pajek/), and following the steps described by Colicchia and Strozzi (2012). The aim of the main path analysis is to identify the most relevant papers about the topic and searching in them for specific data science tools linked to the lean context.

The keywords' co-occurrence network analysis is based on a combined approach including weighted and parameterized modularity-based clusterisation and mapping of bibliometric networks, developed by Waltman et al. (2010). Considering all the keywords (i.e., both author keywords and indexed keywords), the software VoSviewer (http://www.vosviewer.com/) gave as results the keywords' co-occurrence network map (figure 3) and the list of the keywords' clusters items (Table 2). In the keywords' co-occurrence network map, the clusters are identified with different colours and the dimension of the nodes, namely keywords, represents their occurrence. The clusters contain keywords that are used more often together in the same papers, and hence they refer to the same specific research area. Therefore, the analysis of the keywords' clusters can show if literature has some defined combinations of tools of data science and practices or techniques related to lean. As in Ciano et al. (2019), names are given to clusters with the aim to reflect their content and ease their reading and understanding.

Table 1: Applied method

Data	Software	Software Algorithm	Software Output (Tools for the Analysis)
Citations	Pajek	Global key route main path extraction	Publications' main path
Keywords	VoSviewer	VOS mapping combined with modularity-based clustering	Keywords' co-occurrence network map and list of Keywords'

clusters items

4.Publications' main path analysis

Figure 2 represents the main path of the biggest connected component of the citations' network. It consists of nine papers whose publication dates range from 2012 to 2019, demonstrating that the topic has been experiencing a real development only in recent years.

The analysis of the publications' main path reveals that most of its nodes refer to the analysis of data only in a statistical way. Indeed, in all the papers, except the one by Rauch et al. (2016), "data analysis" is present in the abstract or in the keywords to indicate the analysis of the variance among the results from survey researches about lean. Therefore, these papers did not focus on the support of data analysis on lean, but they only adopted data analysis to make inferences about lean.

However, the two papers by Tortorella et al. (2018b; 2019a), even if they use the term "data analysis" as part of "multivariate data analysis", describe the possible advantage of combining lean and data science.

In the first article (Tortorella et al. 2018b), the authors stated that big data can facilitate the calculation and the processing between needs and functions for large data. This allows differentiation of customers' categories and hence of product and solutions, thereby positively affecting customer-related lean practice.

In the second paper (Tortorella et al. 2019a), the authors pointed out that product/service technologies such as cloud service, IoT and big data analysis, nowadays are required for efficient adoption of flow-related lean practice. Moreover, they stated that the availability, processing, and analysis of big data allow the prototyping and integrating design and commissioning approach that can anticipate manufacturing issues.

The paper by Rauch et al. (2016), instead, highlighted that the use of Industry 4.0 technologies, and especially the implementation of ERP-workflows, the adoption of standardized data format such as OPC-UA and of PLM software, can improve data processing leading the lean product development towards a smart version.

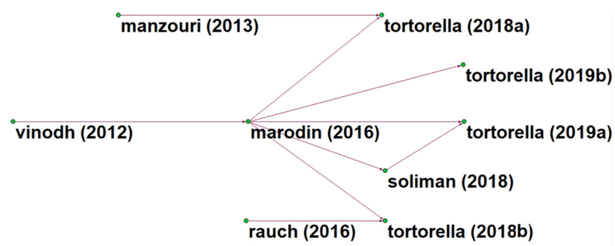

Figure 2: publications' main path.

It includes: Vinodh and Joy (2012); Manzouri et al. (2013); Marodin et al. 2016; Rauch et al. (2016); Soliman et al. (2018); Tortorella et al. (2018a); Tortorella et al. (2018b); Tortorella et al. (2019a); Tortorella et al. (2019b).

5.Keywords' co-occurrence network analysis

Table 2 reports the results of the application of clustering techniques to both indexed and author co-occurred keywords. It is possible to identify 11 clusters of keywords that have been used together by either authors or Scopus for describing the contributions to literature.

The analysis of the keywords constituting the clusters reveals the topics that, in particular, have been discussed by researchers when dealing with data and lean/six sigma. Keywords are shown in order of their links strength (i.e. a proxy of importance) within the cluster. The clusters are automatically ordered based on the number of keywords constituting them.

Below a brief description of main elements of each cluster is provided to better understand how literature has discussed the topic.

Table 2: Clusters of keywords

Cluster	Keywords
1) JIT systems	data reduction, data analysis, statistical methods, optimization, benchmarking, just in time production, mathematical models, computer software, visualization, database systems, software engineering, user interfaces, performance, scheduling, computer architecture, data structures, object oriented programming, statistics, computer simulation, just in time, java programming language, semantics, combustion
2) Lean construction	lean production, project management, data handling, productivity, information systems, production control, construction, lean construction, data analytics, human resource management, construction industry, digital storage, supply chains, construction projects, information analysis
3) Lean product development	data visualization, product development, product design, knowledge management, design, competition, life cycle, industry, management, research, product development process, commerce, systems engineering
4) Continuous improvement in healthcare	continuous improvements, lean, agile manufacturing systems, process improvement, tools, lean principles, exhibitions, health care, hospitals, value stream mapping, continuous improvement, exploratory data analysis, statistical data analysis
5) Six Sigma	big data, work simplification, six sigma, process engineering, process monitoring, data mining, lean six sigma, quality improvement, data science
6) Lean manufacturing	manufacture, surveys, lean manufacturing, design/methodology/approach, industrial research, lean management, automotive industry, manufacturing industries, regression analysis
7) Decision making	decision making, industrial management, societies and institutions, data processing, standards, data gathering, data collection, statistical process control
8) Total productive maintenance	information management, industry 4.0, internet of things, planning, real time systems, automation, maintenance
9) Decision support systems	quality control, customer satisfaction, quality assurance, problem solving, decision support systems, quality function deployment
10) Process control	data acquisition, process control, education, engineering education, students,

11) Total quality management	total quality management, quality management, organizational performance, tqm

5.1 Cluster 1 - Just in time systems

The biggest cluster of the co-occurrence keywords network deals with management in manufacturing systems. In *JIT* systems, focus on order retrieval activities is worth the minimization of production delays and inventory accumulation between production stages (Elsayed et al. 1993). The use of data in previous researches was devoted to the evaluation of the performance of JIT systems and identification of those characteristics that favor its implementation. Nowadays research on JIT systems focuses on how to make them adaptive. Latest research uses big data and *data analysis* to support the *scheduling* step of planning. Moreover, data *visualization*, combined with *semantics*, emerges as a way to explore and gain insights from time series. Within this cluster, *data reduction*, as the transformation of digital information to its meaningful part, emerges a fundamental step of any study that deals with *data analysis*.

5.2 Cluster 2 - Lean construction

Typical *lean production* interest in *production control* has increased over recent years, especially in the *lean construction* context. As technology has been developed to the point where it is possible to remotely locate people, equipment and products in supply chains, and real time *production control* on *construction* site is fundamental, several researches have focused on the definition of the types of the tracking data and analyses towards the increase in *productivity* of *construction projects*. On the other hand, *construction* logistics is an essential part of lean construction for both project management and cost aspects. *Data analysis* of vehicle movements suggested that construction transportation costs can be successfully monitored and managed thanks to the gathering and analyses. Similarly, these concepts can be extended to general *project management*.

5.3 Cluster 3 - Lean product development

Lean product development is the application of lean principles to product development to reduce wastes in the *product development process* and to focus on value adding activities (Rauch et al. 2016). Regarding this specific context, research has concentrated on the role of product *lifecycle* management tools, fundamental to track all data regarding a product and provide a base for future modifications in *product design*. Of course, the use of such a tool, as well as digital mock-ups, smart sensorics, embedded systems involve the collection and management of big data.

5.4 Cluster 4 - Continuous improvement in healthcare

Lean is constituted by two pillars, *continuous improvement* (kaizen) and respect for people, and several *lean principles* (Pakdil and Leonard 2014). Numerous conference papers in the area of lean healthcare presented at *exhibitions* the implementation within *hospital* areas of *tools* supporting *process improvement*. Among them, the mainly applied tools are *Value Stream Mapping*, spaghetti diagrams, Ishikawa charts, and *statistical data analyses* and *exploratory data analyses*.

5.5 Cluster 5 - Six Sigma

The *Six Sigma* approach has evolved to be an extension of quality management (Siddiqui et al. 2016) and works focused on its tools and analyses are classified by SciVal (https://www.scival.com/landing) as "*Six Sigma* | *Work simplification* | Sigma level". Exploratory data analyses and the use of bar charts and box plot to accomplish *process monitoring* and *process engineering* are typical of a *Six Sigma* approach and presented by most researches. Moreover, also sophisticated quality and statistical concepts can be found in research dealing with *Six Sigma* applications, as structured and data-driven approaches. Most recent research is interested in developing frameworks integrating *data science*, *big data* and analytics and the phases of *lean six sigma* in all the dimensions such as volume, variety, velocity and veracity of *big data*.

5.6 Cluster 6 - Lean manufacturing

Lean management in *manufacturing* systems represents the main application studied in *industrial research* works published by the International Journal of Production Research (Ciano et al. 2019). In particular, *automotive industry* is known as the industry where lean production concepts were born and as the one that presents most implementation cases. Numerous studies regarding *lean manufacturing* analyze data from companies collected via *surveys* in order to get some insights. In these researches, particular attention is devoted to the definition of *design/methodology/approach* used.

5.7 Cluster 7 - Decision making

Identification of the wastes affecting the organization and the subsequent selection of value stream mapping tools are known in literature as a complex multi-criteria *decision making* problem (Singh et al., 2006). The use of data has been and is still crucial for continuous improvement programs. For this reason, research has reported cases from both *industrial management* contexts and *societies and institution*. Nowadays research on decision making in lean contexts is studying how the recent opportunities for *data gathering, collection* and *processing* can support *decision making*.

5.8 Cluster 8 - Total productive maintenance

Total productive *maintenance* is one of the aspects of kaizen that seeks for the stability of the process, ensuring availability of resources and increasing overall equipment productivity (Cigolini and Turco, 1997). In a lean context maintenance is twofold: autonomous and planned. *Planning maintenance* activities have always relied on data analyses. Nowadays many researches are exploring the possibilities related to predictive approaches enabled by *real time* data gathering and analysis enabled by *internet of things*-based tools. In particular, *automation* of maintenance processes represents the objective of these research activities.

5.9 Cluster 9 - Decision support systems

Problem solving represents one of the main aspects of *quality assurance* that, together with *quality control*, define TQM and are aimed at *customer satisfaction* (Hellsten and Klefsjö, 2000). In this context, research has focused on the development and the identification of *decision support systems* based on data analysis as tools to achieve of *quality control* and *assurance*.

5.10 Cluster 10 - Process control

Process control is known as potential for cost saving in modern manufacturing systems and six sigma methodologies have exploited this activity for continuous process improvement (Stojanovic and Milenovic, 2018). Recently, *data acquisition* in real time allows for a *process control* method that exploits big data analytics. Several studies have been dedicated to the combination of six sigma principles, such as DMAIC, and big data analytics to study the performance of higher *education* institutions. Other researches deal with how to teach six sigma tools for *process control* in engineering education.

5.11 Cluster 11 - Total quality management

TQM is a holistic quality management approach that considers the entire value chain and stresses the role of human factors. Notwithstanding the great advantages that *TQM* could provide, applying it is hard and intensely affected by critical factors (Hietschold, Reinhardt, and Gurtner 2014). Many researches over years have dealt with the identification of the critical success factors of *quality management* and *TQM* implementations and their impact on primary performance measures. In order to achieve such result, data analyses, such as Partial Least Squares Structural Equation Modeling, have been exploited.

6.Discussion

From the analysis of clusters, it is possible to understand that the use of data in lean/six sigma contexts have been widely applied, ranging from JIT to TQM and TPM as lean principles, and from construction to healthcare or automotive as industries. The adoption of data science in the context of the identified areas of lean principles and six sigma is not surprising, as they all traditionally rely on the use of data, mathemetical and statistical tools for their implementation. As well, the industrial contextes that more adopt the combination of lean and data science are well-known as relying on the use of data.

Moreover, analyzing the clusters structures and papers using them, some keywords emerge as significant within this area of research. First, *data reduction* is mentioned by most of researches as fundamental step of any study that deals with the use of data. *Statistical* and *exploratory data analyses* appear in the cluster involving continuous improvement activities conducted in healthcare contexts. Nonetheless these analyses are typical of six sigma that, due to the aptitude to the use of data, has been recently studied in the development of frameworks for its integration with data science, big data and analytics. *Decision making*, transversal from industry to institutions, emerges as another aspect that could benefit of the evolution of data science. Last, research devoted to *maintenance* represents a knowledge base on the use of *real time* data gathering and analysis.

On the other hand, some concepts emerge as underrepresented, such as jidoka, one of the main elements of lean, aimed at improving the quality of the product and people safety. It seems that research has not focused its interest in the use of data to support this element. On the other hand, the previous research by Ciano et al. (2019) proved that, notwithstanding its importance, jidoka, as well as its synonymous 'autonomation', is rarely mentioned as a keyword. Hence, research on this topic can be overlooked by the adopted methodology.

From the description of the main path, we can tell that the topic has been experiencing a real development only in recent years. Thus, we expect the joint use of data science elements and lean to increase significantly.

According to this study, the expected areas of future research will probably cover underrepresented adoptions of data science activities in lean practices (e.g. jidoka) as well as combined adoptions of data science and lean principles in manufaturing contextes different from healthcare, contstruction and automotive.

7.Conclusions

Due to bibliometric tools the present paper was able to identify what elements of data science have been exploited jointly with lean over years. The applied methodology recognized the main contributions to this topic and highlights that interest is rising in recent years. Moreover, the co-occurrence network analysis helped in identifying the lean areas that have already experienced particular implementations and that seem to be promising.

This work contributes to the body of knowledge, through the identification of the principles and industries that have been mainly interested by the use of data, providing a knowledge base on developed topics and underrepresented ones.

The research finds its main limitation in the materials. In fact, the choice of screening all contributions found in Scopus with no restriction in time may have made old contributions (mainly related to inferences about lean and its principles) within the data set. In fact, the use of data to support the use of a tool or practice is a concept that is more recent and related to technological evolutions.

As future research directions, a systematic review of a more recent data set is suggested to highlight only the cutting edge contributions to the topic, moreover, the development of a framework for the joint use of data science and lean would provide a reference to the effective implementation.

References

Buer, S. V., Strandhagen, J. O., & Chan, F. T. (2018). The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda. International Journal of Production Research, 56(8), 2924-2940.

Ciano, M. P., Pozzi, R., Rossi, T., & Strozzi, F. (2019). How IJPR has addressed 'lean': a literature review using bibliometric tools. International Journal of Production Research, 1-34.

Cigolini, R., & Turco, F. (1997). Total productive maintenance practices: a survey in Italy. Journal of Quality in Maintenance Engineering.

Cleveland, W.S. (2001), "Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics," International Statistical Review, 69, 21–26.

Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort‐Martorell, X., & Reis, M. S. (2016). How can SMEs benefit from big data? Challenges and a path forward. Quality and Reliability Engineering International, 32(6), 2151-2164.

Colicchia, C., & Strozzi, F. (2012). Supply chain risk management: a new methodology for a systematic literature review. Supply Chain Management: An International Journal, 17(4), 403-418.

Donoho, D. (2017). 50 years of data science. Journal of Computational and Graphical Statistics, 26(4), 745-766.

Elsayed, E. A., M. K. Lee, S. Kim, and E. Scherer. 1993. "Sequencing and Batching Procedures for Minimizing Earliness and Tardiness Penalty of Order Retrievals." The International Journal of Production Research 31 (3): 727–738.

Hellsten, U., & Klefsjö, B. (2000). TQM as a management system consisting of values, techniques and tools. The TQM magazine.

Hietschold, N., R. Reinhardt, and S. Gurtner. 2014. "Measuring Critical Success Factors of TQM Implementation Successfully – A Systematic Literature Review." International Journal of Production Research 52 (21): 6254–6272.

Liu, J. S., & Lu, L. Y. Y. (2012). An integrated approach for the main path analysis: The development of the hirsch index as an example. Journal of the American Society for Information Science and Technology, 63(3), 528–542.

Lucio-Arias, D., & Leydesdorff, L. (2008). Main-path analysis and path-dependent transitions in HistCite™-based historiograms. Journal of the American Society for Information Science and Technology, 59(12), 1948-1962.

Manzouri, M., Ab Rahman, M. N., Saibani, N., & Zain, C. R. C. M. (2013). Lean supply chain practices in the Halal food. International Journal of Lean Six Sigma.

Pakdil, F., & Leonard, K. M. (2014). Criteria for a lean organisation: development of a lean assessment tool. International Journal of Production Research, 52(15), 4587-4607.

Pozzi, R., & Strozzi, F. (2018). How assembly systems are adopting the technologies of I40: a preliminary landscape. XXII Summer School "Francesco Turco"–Industrial Systems Engineering, 2018, 369-375.

Rauch, E., Dallasega, P., & Matt, D. T. (2016). The way from lean product development (LPD) to smart product development (SPD). Procedia CIRP, 50, 26-31.

Rejikumar, G., Aswathy Asokan, A., & Sreedharan, V. R. (2020). Impact of data-driven decision-making in Lean Six Sigma: an empirical analysis. Total Quality Management & Business Excellence, 31(3-4), 279-296.

Runkler, T. A. (2012). Data Analytics. *Wiesbaden: Springer. doi*, *10*, 978-3.

Siddiqui, S. Q., F. Ullah, M. J. Thaheem, and H. F. Gabriel. 2016. "Six Sigma in Construction: A Review of Critical Success Factors." International Journal of Lean Six Sigma 7 (2): 171–186.

Singh, R. K., Kumar, S., Choudhury, A. K., & Tiwari, M. K. (2006). Lean tool selection in a die casting unit: a fuzzy-based decision support heuristic. International journal of production research, 44(7), 1399-1429.

Soliman, M., Saurin, T. A., & Anzanello, M. J. (2018). The impacts of lean production on the complexity of socio-technical systems. International Journal of Production Economics, 197, 342-357.

Stojanovic, N., & Milenovic, D. (2018, December). Data-driven Digital Twin approach for process optimization: an industry use case. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 4202-4211). IEEE.

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(22), 6572-6591.

Tortorella, G. L., Fettermann, D., Cauchick Miguel, P. A., & Sawhney, R. (2019b). Learning organisation and lean production: an empirical research on their relationship. International Journal of Production Research, 1-17.

Tortorella, G. L., Giglio, R., & Limon-Romero, J. (2018a). Supply chain performance: how lean practices efficiently drive improvements. Journal of Manufacturing Technology Management.

Tortorella, G. L., Giglio, R., & van Dun, D. (2019a). Industry 4.0 adoption as a moderator of the impact of lean production practices on operational performance improvement. International journal of operations & production management.

Tortorella, G. L., Miorando, R. F., Fries, C. E., & Vergara, A. M. C. (2018b). On the relationship between Lean Supply Chain Management and performance improvement by adopting Industry 4.0 technologies.

Vinodh, S., & Joy, D. (2012). Structural equation modelling of lean manufacturing practices. International Journal of Production Research, 50(6), 1598-1607.

Waltman, L., van Eck, N. J., & Noyons, E. C. (2010). A unified approach to mapping and clustering of bibliometric networks. Journal of Informetrics, 4(4), 629-635.

Womack, J. P., Jones, D. T., & Roos, D. (2007). The machine that changed the world: The story of lean production--Toyota's secret weapon in the global car wars that is now revolutionizing world industry. Simon and Schuster.

Zhong, R. Y., Xu, C., Chen, C., & Huang, G. Q. (2017). Big data analytics for physical internet-based intelligent manufacturing shop floors. International journal of production research, 55(9), 2610-2621.