International Journal of Production Research

ISSN: 0020-7543 (Print) 1366-588X (Online) Journal homepage: https://www.tandfonline.com/loi/tprs20

From data to big data in production research: the past and future trends

Yong-Hong Kuo & Andrew Kusiak

International Journal of Production Research, 2019
Vol. 57, Nos. 15–16, 4828–4853, https://doi.org/10.1080/00207543.2018.1443230

From data to big data in production research: the past and future trends

Yong-Hong Kuo[a] and Andrew Kusiak[b]*

[a]*Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong;* [b]*Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, IA, USA*

(*Received 13 October 2017; accepted 16 January 2018*)

Data have been utilised in production research in meaningful ways for decades. Recent years have offered data in larger volumes and improved quality collected from diverse sources. The state-of-the-art data research in production and the emerging methodologies are discussed. The review of the literature suggests that production research enabled by data has shifted from that based on analytical models to data-driven. Manufacturing and data envelopment analysis have been the most popular application areas of data-driven methodologies. The research published to date indicates that data mining is becoming a dominant methodology in production research. Future trends and opportunities for data-driven production research are presented.

Keywords: production; data; data mining; data-driven models; big data; smart manufacturing; data envelopment analysis; simulation

1. Introduction

Data have been used in production research since its infancy. For example, the very first paper included in the inaugural issue of the *International Journal of Production Research* (Kendrick 1961) discussed prediction of service needs for automotive components based on the historical component return data. Naturally, in today's standards the data-sets used for research decades ago were small due to the limited data collection and storage ability (Davenport and Dyché 2013). The methodologies to conduct research did not face much difficulty in processing the data as the volumes were generally limited. The main challenges facing the use of data in production research in the early times were in data scarcity and time-lapse (Ngai et al. 2017).

The emergence of the computing and storage technology has positively impacted the quality and quantity of the data. The modern era data are stored in relational databases (Hoffer 2011), supervisory control and data acquisition (SCADA) systems (Boyer 2009) and data warehouses using cloud computing as one of the preferred solutions (Buyya et al. 2009). The wider variety and greater volumes of data allow researchers and practitioners to develop better understanding of the issues facing production systems, forecast demands, monitor and control production processes, optimise production decisions and derive insights into production management.

Owing to the advanced information technologies, the data collected by enterprises can be enormous, often referred to as *big data*. Big data not only is about the *volume* (Kitchin 2014), but also has other characteristics such as *variety*, *velocity* and *veracity* (Jin et al. 2015). According to the McKinsey Global Institute report (Manyika et al. 2011), big data implies volumes exceeding the capability traditional database software tools to capture, store, manage and analyse.

While data have been essential for production research, it appears that comprehensive studies of its use and importance are scarce. Measuring the use and significance of data in production research is not easy. A meaningful research project involves a thorough analysis of the research landscape. For the purpose of this paper, the data extracted from the journals included the Taylor & Francis Online library are considered to reflect its use of data in various production research topics (see in Table 1).

All topics in Table 1, but topic 1, have been selected based on the scope of the *International Journal of Production Research (IJPR)*. Topic 1 'data' were used as a reference to the remaining 25 production topics. The Taylor & Francis Online digital library was searched on the keywords included in each topic to provide the entries in the third column of Table 1. Each entry is the number of published papers that include the corresponding keywords. The topics in Table 1 are sorted in the descending order of the number of published papers. It has been determined that if the keyword 'data'

*Corresponding author. Email: andrew-kusiak@uiowa.edu

Table 1. The number of papers published at Taylor & Francis Online associated with the selected production research topics.

Topic no.	Topic	Number of published papers
1.	Data	2,008,014
2.	Production data	1,228,892
3.	Product development data	1,101,381
4.	Production analysis data	1,068,781
5.	Process control data	945,720
6.	Process modelling data	925,045
7.	Human factors data	610,570
8.	Quality control data	595,171
9.	Group technology data	583,236
10.	Production planning data	522,807
11.	Facility location data	236,592
12.	Design for manufacturing data	224,673
13.	Production scheduling data	148,389
14.	Process planning data	145,518
15.	Supply chain data	95,081
16.	Sustainable manufacturing data	91,546
17.	Green manufacturing data	58,021
18.	Manufacturing automation data	48,884
19.	Industry 4.0 data	40,356
20.	Intelligent manufacturing data	40,338
21.	Lean manufacturing data	19,236
22.	Facility layout data	17,330
23.	Cloud manufacturing data	12,982
24.	Smart manufacturing data	11,775
25.	Process modularity data	3483
26.	Product modularity data	3418

would be ignored, the number of the published papers would be larger, however, the order of the topics would remain the same.

What Can Be Learned From the Data in Table 1?

The number of published papers in Table 1 indicates the magnitude of references made to the keyword ‘data’ in the context of the corresponding topic. In no way this number reflects importance of any topic, rather it reflects its coverage in the journals captured in the Taylor & Francis Online library. The topics have the following characteristics:

(a) They may vary in the granularity and features of the library,
(b) They could overlap,
(c) A topic with a higher number of papers may include less prominent topics.

It may be reasonable to assume that the importance and the volume of data in the domain associated with a topic may be positively correlated to the number of papers published. The ‘production’ topic (No. 2) has the largest coverage measured in the number of published papers and therefore it is selected to review the past developments and trends in data research and applications.

Thus, the goals of this review paper are to:

- identify the trends of production data research;
- review the past production data research; and
- discuss and explore future opportunities.

The organisation of the paper is as follows. Section 2 analyses the trends in the usage of data in production research. Section 3 reviews the past developments of production research enabled by data. The most recent developments of production data research are reviewed in Section 4. Section 5 offers discussion of the literature and future research directions. Section 6 concludes this paper.

2. Trends in the usage of data in production research

The goal of this section is to review the papers published in *IJPR* to gain insights into the trends in the usage of data in production research. The number of papers published in *IJPR* (as of 29 September 2017) since its inception with the word ‘data’ in the title or as a keyword is plotted in Figure 1. In total, 289 papers have been identified. Note that keywords of the *IJPR* papers were introduced in 2004, which could potentially underestimate the trend if the title would not reflect the data use. Nevertheless, Figure 1 provides insights into the usage of data in production research. Before 1990, production research based on data was rather sporadic; it became a regular, but not a major topic, in the period from 1990 to 2004; and it intensified since 2005. In the recent years, it has become a widely researched topic. The years 2006, 2010 and 2017 had experienced significantly more papers published relevant to data. This could be due to the special issues focused on the topics related to data research: ‘Data mining and applications in engineering design, manufacturing and logistics’ (Feng and Kusiak 2006), ‘RFID technology and applications in production and supply chain management’ (Ngai 2010) and ‘Using big data to make better decisions in the digital economy’ (Tan et al. 2017).

To identify the most popular topics on data research in production, tag clouds from the keywords of 289 papers for four different periods – 2005–2009, 2010–2013, 2014–2016 and from 2017 – are presented in Figures 2–5. These time intervals were chosen in a way that: (i) the keywords were available in the corresponding years and (ii) the number of papers published with the word ‘data’ in the title or as keywords was about equal. The size of each word in Figures 2–5 is directly proportional to its frequency. Note the following: (i) the word ‘data’ was removed from the tag clouds as it was popular in all papers of focus, (ii) the common words (‘a’, ‘the’, ‘of’, ‘for’, etc.) have been removed and (iii) stemming to group words originating from the same ‘word’ (e.g. ‘manufacture’ and ‘manufacturing’ were grouped to the same word) was performed. The following observations are made based on Figures 2–5:

(1) ‘Manufacture/manufacturing’, ‘analysis’, ‘envelopment’, ‘mining’, ‘process’, ‘system’, ‘product’, ‘model’, ‘design’ and ‘management’ have been consistently used as keywords since 2005;
(2) The word ‘big’ has emerged in the recent years, largely in the context of ‘big data’;
(3) Data from online platforms and social media (as reflected from the words ‘online’ and ‘social’) have become emerging forms of data in production research;
(4) ‘Prediction’ has become an emerging topic in the recent data research.
(5) The interests in ‘neural network’ research has grown recently.

To better understand the trends in production data research, the relative frequency of each topic, defined as the number of papers with the keyword topic divided by the total number of papers published in the period, is computed. The relative frequency of the topics ‘mining’, ‘process’, ‘system’, ‘product’, ‘model’, ‘design’ and ‘management’ were quite stable during the four periods, suggesting their consistent popularity. Figure 6 presents the relative frequency of selected topics that exhibit interesting trends:

(1) ‘Manufacturing’ was observed to lose popularity from 2005 to 2016, but has attracted more attention in 2017;
(2) The word ‘control’ has become less popular over the years;

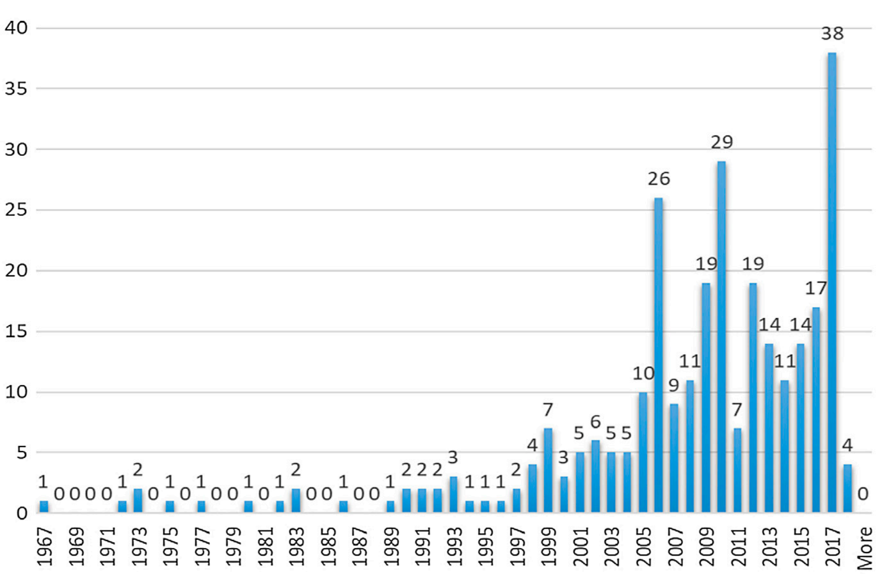

Figure 1. Number of papers with the word ‘data’ in the title or as a keyword published in *IJPR* (as of 29 September 2017).

Figure 2. Tag cloud of keywords from *IJPR* papers published in years 2005–2009.

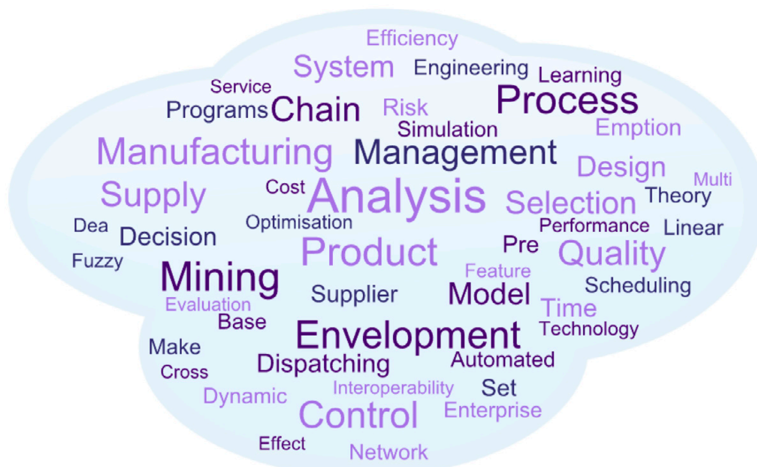

Figure 3. Tag cloud of keywords from *IJPR* papers published in years 2010–2013.

(3) 'Envelopment' or 'DEA' (data envelopment analysis) was a consistently used keyword until 2016, but became less popular starting from 2017;
(4) While 'analysis' is still a keyword, it has become less popular from 2017. Meanwhile, the word 'analytics' (which is often interpreted as the use of advanced mathematical, statistical or computational methods for the discovery of knowledge or insights from large data-sets) has become an emerging keyword in the recent years.

These observations offer insights into the trends in production data research over the past 12 years:

(1) The general goal of the research conducted over the 12-year period was consistent – to analyse and improve the 'process', 'system', 'product', 'design' and 'management' in the context of 'manufacturing';
(2) 'Modelling' research was consistently popular, but the models have shifted from analytical and less data-intensive (e.g. DEA) to more data-intensive approaches (e.g. data mining);
(3) The goal of data research in recent years had transformed from more focused on 'control' to the discovery of meaningful insights (i.e. analytics) from large volumes of data (i.e. big data). The data research will likely result in more accurate prediction models;
(4) Other forms of data became available (e.g. from online platforms and social media) due to the advancement of technologies and could be incorporated in production research.

Figure 4. Tag cloud of keywords from *IJPR* papers published in years 2014–2016.

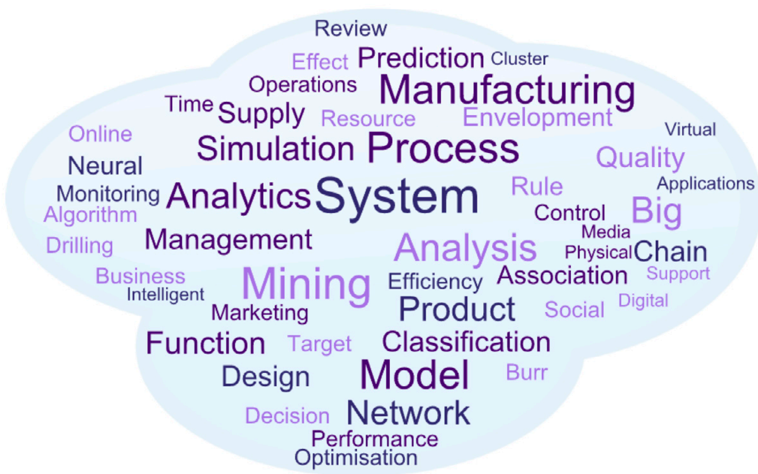

Figure 5. Tag cloud of keywords from *IJPR* papers published since 2017.

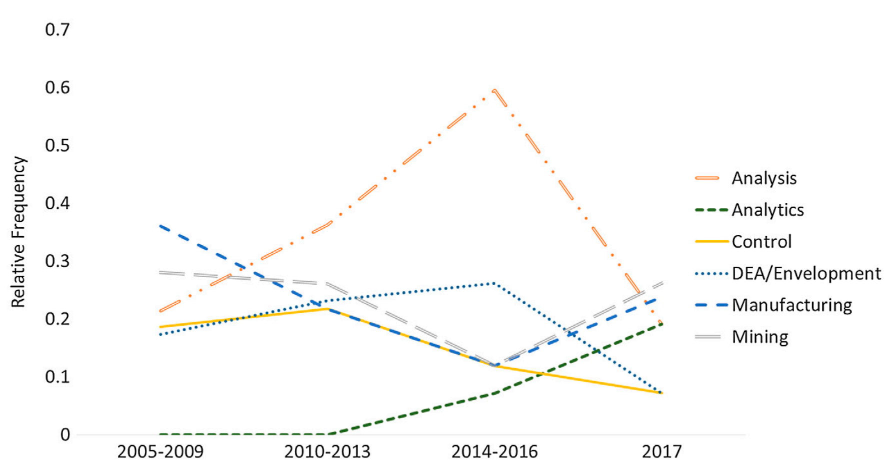

Figure 6. Relative frequency of the production data research topics.

The above observations and insights have provided motivation to review the past developments in production data research and to explore the future opportunities in this domain.

3. Past developments

In this section, the past developments (the period between 1961 and 2013) in the usage of data in production research are discussed. As indicated in Figure 1, the usage of data in production research has drawn more attention since 2000. Thus, the introduction and growth stages are defined as the periods between 1961 and 1999, and between 2000 and 2013.

3.1 *Introduction stage: the 1961–1999 period*

At the introduction stage, methodologies developed and applied in production research were largely analytical. Due to the limited capability to collect and process data, the usage of data was often not a primary focus but rather to estimate parameters of the analytical models. Since the variety and volume of data in the period were manageable, the focus of research was mainly to develop the theory to demonstrate the effectiveness of the methodologies. The computational complexity implied by the data was not an issue.

The review of the past developments of production data research begins with the very first paper by Aberg (1967), among the 289 papers identified in the title and keyword search. The author studied the sampling and observational errors of work element time and conducted statistical analysis of the data. In the 1970s, papers on exponential smoothing in process control (e.g. Wortham 1972), estimation of learning curves (e.g. Towill 1973) and exponential smoothing on learning curve data (e.g. Towill 1977) were published. These papers mainly dealt with the time series data that were not intensive. At the very beginning of the introduction stage, development of statistical and forecasting methods was the focus.

In the 1980s, data utilisation in production research was still rare. Notable applications include the generation of data at a numerically controlled machine tool for the design of free-form surfaces (e.g. Broomhead and Edkins 1986) and the groupability of a data-set for the adoption of group technology (e.g. Chandrasekharan and Rajagopalan 1989). The utilisation of data was focused on assisting in product design and enhancing manufacturing.

In the 1990s, production research with the focus on data was regularly conducted and the applications became more diverse. One of the most popular aspects involving the use of data was group technology. Some of the examples were machine–component grouping (e.g. Kusiak 1987; Gupta and Seifoddini 1990), development of quantifiable measures for grouping with fuzzy features (e.g. Ben-Arieh and Triantaphyllou 1992), production sequence data for cell formation (e.g. Nair and Narendran 1998) and cell formation with the use of ordinal and ratio-level data (e.g. Nair 1999). In these applications, the data was used to compute similarity measures for grouping decisions. Clustering was a widely applied solution approach in group technology. Another aspect of data research during this period was performance monitoring or control of processes (e.g. Kim and Kolarik 1992; Kanagawa, Tamaki, and Ohta 1993; Hwarng 1995; Shore 1998). Data were deployed to construct control charts in most of these applications. Papers on the enhancement of manufacturing systems were published, e.g. the use of data in support of flexible manufacturing systems (e.g. O'Keefe and Haddock 1991) and identification of best operating practices in cellular manufacturing systems (e.g. Talluri, Huq, and Pinney 1997). Shafaei and Brunn (1999a, 1999b) addressed the issue of inaccurate data in a job shop environment and proposed scheduling rules to improve the system performance. Research on construction of surfaces with coordinate measuring data in production processes (e.g. Lee, Chen, and Lin 1990; Chiang and Chen 1999) is also noted.

In summary, at the introduction stage, the domains of focus included: product design, operational efficiency enhancement and process control. The data usage was rather standard, e.g. estimation of parameters of analytical models and as input for statistical models. The volume and variety of data utilised were limited due to a small number of variables involved. The primary goal of the research was to demonstrate the effectiveness of the solution methodologies. The data utilised in the research came from industry or were imulated. Due to the limited variety of data and the data sources, the data-sets were well-organised and were easy to manage. Research challenges did not arise from data processing but the methodologies.

3.2 *Growth stage: the 2000–2013 period*

Since 2000, production research enabled by data was expanding. Table 2 provides the list of papers published in *IJPR* from 2000 to 2013 with the word 'data' in the title or as keyword and over 30 citations (according to CrossRef citations, as of 29 September 2017). In total, 19 papers were retrieved. The number of citations can be regarded as an indicator of the popularity of the research and its impact.

Table 2. The papers published in *IJPR* in the period 2000–2013 with the word ‘data’ in the title or as a keyword with over 30 citations (according to CrossRef citations).

Paper title	Year	Authors	Citations[a]
Enterprise risk management: a DEA VaR approach in vendor selection	2010	D.D. Wu and D. Olson	102
An application of data envelopment analytic hierarchy process for supplier selection: a case study of BEKO in Turkey	2007	M. Sevkli, S.C.L. Koh, S. Zaim, M. Demirbag and Ekrem Tatoglu	93
Data mining-based methodology for the design of product families	2004	B. Agard and A. Kusiak	73
A leanness measure of manufacturing systems for quantifying impacts of lean initiatives	2008	H.-D. Wan and F.F Chen	64
From closed-loop to sustainable supply chains: the WEEE case	2010	J. Quariguasi Frota Neto, G. Walther, J. Bloemhof, J.A.E.E van Nunen and T. Spengler	63
Integrating data mining and rough set for customer group-based discovery of product configuration rules	2006	X.-Y Shao, Z.-H Wang, P.-G Li and C.-X. J Feng	53
Robust closed-loop supply chain network design for perishable goods in agile manufacturing under uncertainty	2012	A. Hasani, S.H. Zegordi and E. Nikbakhsh	51
Process monitoring for multiple count data using generalized linear model-based control charts	2003	K.R. Skinner, D.C. Montgomery and G.C. Runger	49
Assessing computer numerical control machines using data envelopment analysis	2002	S. Sun	48
Supplier evaluation and selection: an augmented DEA approach	2009	T. Wu and J. Blackhurst	48
Supplier selection using analytic network process and data envelopment analysis	2012	R.J. Kuo and Y.J. Lin	47
Data driven bottleneck detection of manufacturing systems	2009	L. Li, Q. Chang and J. Ni	45
Data mining: manufacturing and service applications	2006	A. Kusiak	43
Wavelet-based SPC procedure for complicated functional data	2006	M.K. Jeong, J.-C. Lu, and N. Wang	43
Application support to product variety management	2008	C. Forza and F. Salvador	36
RFID opportunity analysis for leaner manufacturing	2010	A. Brintrup, D. Ranasinghe, and D. McFarlane	36
A comparison of stochastic dominance and stochastic DEA for vendor evaluation	2008	D. Wu and D.L. Olson	35
Optimising product configurations with a data mining approach	2009	Z. Song and A. Kusiak	35
Efficient algorithm for cell formation with sequence data, machine replications and alternative process routings	2004	S. Jayaswal and G.K. Adil	34

[a]The number of citations was obtained from CrossRef citations as of 29 September 2017.

3.2.1 *Data envelopment analysis*

Data envelopment analysis (DEA), aiming at the measurement of product efficiency with empirical data, appears to be one of the most widely researched topics. As observed in Table 2, 7 out of these 19 papers were relevant to DEA. The paper by Wu and Olson (2010) has been most highly cited among these 19 papers. The authors proposed the concept of value-at-risk (VaR) in vendor selection. Their approach aimed at addressing the enterprise risk management. For further discussion on the paper, the reader is referred to Wei and Wang (2011) and a response from the authors (Wu and Olson 2011). In their earlier paper, Wu and Olson (2008) had studied a vendor selection problem where the estimated measures were not precise. Sevkli et al. (2007) applied an envelopment analytic hierarchy process methodology to improve decisions in supplier selection. Wang, Chin, and Leung (2009) addressed some issues of the DEA methodology proposed by Sevkli et al. (2007). Wu and Blackhurst (2009) proposed an augmented DEA for evaluating and ranking suppliers. Kuo and Lin (2012) developed an analysis network process and an envelopment analysis approach to select suppliers, with the consideration of environmental factors. Other papers that adopt DEA for vendor or supplier selection include Talluri, Narasimhan, and Viswanathan (2007), Dotoli and Falagario (2012), Zhang, Lee, and Chen (2012) and Parthiban, Zubar, and Katakar (2013). The number of the papers published and the citations of the papers on vendor or supplier selection using DEA suggest that the data use in this type of research was rather wide.

The remaining two of the seven papers on DEA in Table 2 studied other aspects of production. Wan and Frank Chen (2008), developed a leanness metric based on DEA to quantify the impact of lean initiatives on process improvement in manufacturing systems. Sun (2002) applied DEA to evaluate computer numerically controlled machines and identify a homogeneous set of good systems. Other applications of DEA include:

- evaluation of environmental and economic performances of small- and medium-sized enterprises (Sarkis and Dijkshoorn 2007);
- performance ranking of firms (Chang and Chen 2008);
- assessments of allocation of resources (Chang, Kuo, and Chen 2008);
- production performance assessment (Franklin Liu and Cheng Liu 2008; Bayraktar et al. 2010; Wang, Ho, and Oh 2010);
- performance evaluation of entities in supply chains (Li and Dai 2009; Lu and Hung 2010);
- technology selection (Wang and Chin 2009; Foroughi 2012);
- effect assessment of corporate social responsibility on firms' performance (Lu, Wang, and Lee 2013); and
- maintenance planning (Azadeh et al. 2013).

The primary applications of DEA were relevant to performance assessment with the ultimate goal to support decision-making in the selection process. DEA gained more maturity in this time period, with its early applications in production research found in the 1990s (e.g. Talluri, Huq, and Pinney 1997). Thus, DEA was one of the most popular topics of data research in production, as indicated by the number of papers published and their citations.

3.2.2 *Data-driven algorithms*

Fuelled by advances in information and communication technologies (ICT) around 2000, wider variety and greater volumes of data could be collected and managed. The structure and format of data-sets became more complex and heterogeneous. Traditional statistical approaches faced the challenge of high dimensionality and interactions among the variables. Analytical models, which were developed based on prior knowledge of the systems and underlying phenomena, were not well suited to represent the relationship between many variables. The data-sets consisting of large numbers of measurements enabled researchers to revisit the use of the detailed data to benefit production research.

Data-driven approaches allow discovery of non-trivial patterns from data. Data mining algorithms are the primary model and knowledge discovery tools. Their developments and usage in production inexperience the growth stage. The growing trend of applications of data mining in production research can be illustrated with the *IJPR* special issue 'Data Mining and Applications in Engineering Design, Manufacturing and Logistics', edited by Feng and Kusiak (2006).

The data-driven approaches followed data envelopment analysis (DEA) as the second most popular topic, listed in Table 2. Among the papers in Table 2, 5 of 19 papers were relevant to data mining (if DEA is not considered as one of them). Manikas et al. (2017) analysed research methods adopted by the authors of the papers published in *IJPR* and concluded that data mining methods have gained higher popularity over time in production research. The applications of data mining approaches in production research are diverse. Agard and Kusiak (2004) proposed a two-stage approach with data mining techniques for the design of product families, where the first stage was for customer segmentation and the second stage was for identification of the source of the requirement's variability. Shao et al. (2006) integrated data mining tools, more specifically, clustering and association rule mining, with variable precision rough set to link customers with sets of product specifications and associate the specifications with configuration alternatives. They used a case study of an electrical bicycle to demonstrate the performance of their methodology. Li, Chang, and Ni (2009) proposed a data-driven approach utilising information about production line blockage, starvation probabilities and buffer content records, to identify bottlenecks in manufacturing systems. In a later paper, Li (2009) extended the research to a more complex manufacturing setting. Kusiak (2006) proposed a new data-driven framework to leverage knowledge for decision-making in manufacturing and service systems. He introduced the basic concepts of machine learning and data mining in production research. He also discussed data-driven industrial, medical and pharmaceutical applications. Song and Kusiak (2009) applied a data mining approach to address the issues of product diversity and complexity. They utilised historical sales data to derive rules to determine sub-assemblies and product configurations. Other applications that have adopted data mining in production research include:

- enterprise modelling based on knowledge discovery (Neaga and Harding 2005);
- reduction of risk of having defective products by association rules (Da Cunha, Agard, and Kusiak 2006);
- condition-based maintenance by data fusion techniques (Raheja et al. 2006) or competitive learning and hidden Markov Model (Chinnam and Baruah 2009);
- design of activity knowledge acquisition by a function-based design operation-mining algorithm (Jin and Ishino 2006);
- discovery of relationships between product attributes and causes of failure with an association rule generation algorithm (Buddhakulsomsiri et al. 2006);

- identification of defects or flaws by image processing based on data mining (Dengiz, Smith, and Nettleship 2006; Noh et al. 2010) and integrated association rule and rough sets (Wang, Wang, and Lee 2006);
- technology mining (Hou and Yang 2006);
- retrieval of information about manufacturing processes using a rough set approach (Huang et al. 2006);
- web mining for web-based design and manufacturing (Caramia and Felici 2006);
- prediction of preferred suppliers by rough set theory, support vector machines and feature selection (Tseng et al. 2006);
- detection of changes in distributions of process characteristics (Li, Runger, and Tuv 2006);
- detection of churn activities by a functional mixture model (Qian, Jiang, and Tsui 2006) or with clustering and fuzzy algorithms (Abbasimehr, Setak, and Soroor 2013);
- manufacturing control by actionable data mining (Braha, Elovici, and Last 2007);
- determination of parameters of multiple input multiple output processes with neural networks (Wong, Su, and Hsieh 2007);
- manufacturing cell formation by uncertain association rule mining (Liu et al. 2009);
- determination of dispatching decisions by neural networks (Shiue 2009), decision tree (Metan, Sabuncuoglu, and Pierreval 2010) or machine learning (Guh, Shiue, and Tseng 2011);
- house of quality analysis based on a rough set algorithm (Li and Wang 2010);
- estimation of circuit board assembly times with logic programming (Vainio et al. 2010);
- supplier base management based on rough set theory (Parmar et al. 2010);
- optimisation of manufacturing processes by a rule induction method (Kwak, Kim, and Lee 2010);
- module selection by pattern extraction and entropy maximisation (Da Cunha, Agard, and Kusiak 2010);
- selection of manufacturing technologies with a fuzzy decision tree (Evans et al. 2012); and
- workforce scheduling using association rules (Kim and Nembhard 2013).

Based on the published literature, neural networks, association rule algorithms, clustering techniques, rough set theory, feature selection were the most popular approaches to tackling the production problems. Different from the conventional production data research, data-driven approaches have the capability to process large volumes of data.

3.2.3 *Control charts*

Two of the papers in Table 2 studied the construction of control charts. Skinner, Montgomery, and Runger (2003) proposed a deviance residual-based monitoring scheme to detect a shift in the mean of Poisson counts. They showed that their methodology outperformed a C chart. Jeong, Lu, and Wang (2006) proposed a wavelet-based statistical process control procedure for detecting abnormal events using functional data in manufacturing processes. The adoption of data-driven techniques as a part of research is noted. For example, Chien, Hsu, and Chen (2013) integrated spatial statistics and neural networks to detect wafer bin map defect patterns. Other developments of control charts include:

- design of X charts (e.g. Chou, Chen, and Liu 2001);
- construction of control charts from linguistic data (e.g. Yu, Low, and Cheng 2003);
- estimation of out-of-control rates due to assignable causes process improvement (e.g. Bischak and Silver 2004);
- special cause control charts (e.g. Wang 2005);
- semiconductor machinery control (e.g. Chen et al. 2005);
- exponentially weighted moving average chart with transformed exponential data (e.g. Liu et al. 2007); and
- comparison of control charts for functional data (e.g. Colosimo and Pacella 2010).

3.2.4 *Miscellaneous applications*

The remaining of the papers of Table 2 addressed multiple aspects of production research. Forza and Salvador (2008) discussed the importance of the integration of product configurations, product data management and customer relationship management systems and proposed sales and technical configuration models for product configuration systems. Brintrup, Ranasinghe, and McFarlane (2010) used the seven Toyota manufacturing wastes to provide a toolset in support of RFID establishing a lean manufacturing environment. While the remaining three papers utilised data in their studies, data were not the main focus. Quariguasi Frota Quariguasi Frota Neto et al. (2010) proposed optimisation models to ensure closed-loop supply chains remained sustainable. The environmental impact databases and life cycle data were utilised in their study. Hasani, Zegordi, and Nikbakhsh (2012) considered a closed-loop supply chain network design problem with uncertain interval data. They proposed an interval robust optimisation approach to tackling the problem. Jayaswal and Adil (2004) presented a problem of cell formation in manufacturing with the consideration of sequence data, machine replications and alternative process routings. They proposed a hybrid algorithm of simulated annealing and local search heuristics to tackle the problem.

In summary, while DEA was the most popular approach, data-driven techniques had become an emerging topic, and their applications began rapidly growing since 2000. The techniques and applications of data-driven approaches in production research were diverse, as reflected by the papers reviewed in this session.

4. Recent developments

The goal of this section is to review the recent developments (the period from 2014 to present) of production data research and to classify the research into main areas. Discussion and future directions are provided in Section 5. Based on the literature review, the data research conducted in production has been classified into three major areas: manufacturing, supply chain management and customer research.

4.1 *Manufacturing*

In Section 2, manufacturing has been identified as one of the major research topics involving data. This is supported by the large number of papers published in this area. The recent years have seen significant progress in the capture of process data. Manufacturing processes have become more effective and responsive, and smarter. *Smart manufacturing* is emerging as a new research topic, and its success is closely tied to big data (Kusiak 2017, 2018). Modelling manufacturing processes with the utilisation of data can 'save money, energy and materials' (Kusiak 2017). For an overview of smart manufacturing, the reader may consider the papers by Kang et al. (2016), Kusiak (2017, 2018), Mittal et al. (2017). The recent developments on data-enabled manufacturing applications are presented next.

4.1.1 *Assessing performances of manufacturing systems*

4.1.1.1 *Simulation modelling.* Simulation is broadly used to assess performance of manufacturing systems. Historical data could be utilised to model manufacturing processes and activities. Managers and planners use simulation models to examine different 'what-if' scenarios in decision-making and derive managerial insights. Kacar and Uzsoy (2014) developed a simulation model of a wafer fabrication facility to generate data for the assessment of the performance of product-based and load-based clearing functions (CFs). They found that load-based CFs outperformed product-based CFs. Johansson et al. (2015) developed a discrete event simulation (DES) model that considered NC machine power consumption data and conducted an analysis of data-sets from three factories. Their results suggested that non-value-added activities increase energy consumption and a discrete event simulation approach can help reduce the energy use. Azadeh, Nazari, and Charkhand (2015) proposed a simulation–stochastic DEA approach with the use of historical operational data to design the layout of a job shop facility. Mousavi and Siervo (2017) adopted a DES approach to estimate the key performance indicators (KPIs) of processes from real-time data collected from a control system. Their approach could assist managers in prediction of outcomes of what-if scenarios. Jain, Shao, and Shin (2017) presented the virtual factory concept supported by simulation and data analytics tools. They applied these techniques in generation of data and validation of the results.

Simulation models allow decision-makers to predict outcomes and generate data for scenarios that have not been previously realised. Most of the research on simulation modelling in manufacturing aims to evaluate performance of alternatives and examine benefits and trade-offs of alternatives. However, a shortcoming of simulation modelling is that the model developer must be familiar with the system.

4.1.1.2 *Utilisation of small data.* In practice, manufacturing organisations may not have sufficient data for analysis to support decision-making. It is common that limited information is available for new production processes. To address this issue, Li et al. (2015) utilised small data-sets captured at initial stages of manufacturing and applied a data diffusion method to estimate the data ranges. Following this line of research, Li, Wen, and Chen (2016) applied the small Johnson Data Transformation method to learn from scarce raw data to generate virtual samples.

4.1.1.3 *Manufacturing strategies.* Data provides an objective way to assess performance, derive insights and support research findings. Wang, Törngren, and Onori (2015) proposed a dynamic DEA approach to examine performance of electronic firms. They found that managers were more likely to realise asset impairments when poor performance was reported. Wang and Chien (2016) used balanced scorecard and DEA to assess the performance of LED companies. Godinho Filho et al. (2017) found that companies that have adopted a quick response manufacturing (QRM) approach did not have as much understanding of QRM as lean manufacturing. They also reported that focusing on productivity, low cost and due date delivery can be an obstacle to acquiring the knowledge of QRM. Chakravorty and Hales (2017) conducted a case study in aircraft manufacturing and distribution operations and analysed the process improvement efforts. They found that the improvements could be sustained by applying the experiential learning model with the classical lean and six sigma tools. Lee, Swink, and Pandejpong (2017) studied the relationship between maturity of technologies, team diversity and technical success based on the data from 183 manufacturing process innovation projects. They found that technology maturity can moderate the relationship between some of the factors and the technical success.

4.1.1.4 *Miscellaneous data-driven approaches.* Yu and Matta (2016) proposed a data-driven approach to identify bottleneck machines. They integrated multiple statistical tools to decrease the false detection rate. Omar et al. (2017) developed a data-driven aggregated model as an alternative to the allocated clearing function formulation. Their approach estimated system throughput at discrete work-in-process points and the estimates were used to release production. Zhong et al. (2017) proposed an intelligent manufacturing shop floor framework enabled by the Internet of Things (IoT) and RFID. Their platform captured large volumes of data for analysis and managerial implications. Ye (2017) proposed a reverse engineering approach to mine system data to derive causal relationships in a system with unknown structure.

4.1.2 *Process and quality monitoring*

Big data has been acknowledged as one of the key enablers to process and quality monitoring (He and Wang 2017). The major focus of process and quality monitoring is to establish new statistical foundations of identifying abnormal measurements from samples and develop computational algorithms for effective detection.

4.1.2.1 *Statistical monitoring.* The research on statistical monitoring aims to develop statistical theory and effective monitoring schemes. This line of research has been investigated by researchers for decades. Some of the research in this area performed in 1970s was discussed in Section 3. The high dimensionality and heterogeneous data should be addressed in the monitoring schemes. Bao, Wang, and Jin (2014) considered the thickness variation of wafers at macro- and micro-levels for process monitoring, fault diagnosis and run-to-run control. Grasso, Colosimo, and Pacella (2014) developed a multi-way principal component analysis method for the data dimensionality reduction and sensor data fusion. Their methodology was applied to detect faults impacting the most critical machine components. Riegler et al. (2015) studied the problem of accurate time alignment of independent variables in regression analysis. To overcome the shortcomings of static time alignment, they calculated the time lag dynamically. Their approach resulted in a higher prediction accuracy of final product strength properties. Das et al. (2016) proposed a monitoring approach based on the Poisson–multivariate Gaussian mixed model to address the issues where individual variables are over-spread. Ding, Tsung, and Li (2016) considered the simultaneous presence of continuous and categorical data for monitoring purposes. They proposed a directional exponentially weighted moving average control scheme to address the issue. He et al. (2016) proposed a multivariate generalized likelihood ratio control chart for fault detection from images. Gunaratne et al. (2017) integrated parallelized Monte Carlo simulation into multivariate exponentially weighted mean squared deviation and variance charts to address the issue that processes with high-dimensional variables could not be monitored with reasonable computational efforts. Grasso, Colosimo, and Tsung (2017) studied the multimode process and proposed a methodology based on coupling curve classification and monitoring to automatically switch to the right control chart when new data are fed in.

4.1.2.2 *Process capability.* Process capability indices (PCIs) are used to indicate how well a company performs. Liao (2015) studied the process incapability data samples collected from multiple batches. The author proposed a new method to estimate the true value of the index and reported that the performance of the estimation depended on the degree of process departure. Chen, Wang, and Chang (2017) derived new process capability indices based on the Boole's inequality and DeMorgan's theorem. They used the indium tin oxide film and five-way pipe products to demonstrate the practicality of their results.

4.1.2.3 *Miscellaneous monitoring tools.* Some monitoring tools integrated statistical monitoring schemes and data mining techniques. Some of them adopted big data frameworks such as MapReduce and cloud computing. Jiang et al. (2014) integrated independent component analysis (ICA) and support vector machine data description (SVDD) for monitoring purposes. Their method outperformed a conventional ICA approach. Chen et al. (2016) integrated ICA, SVDD and Durbin–Watson criterion to detect faults. Joo and Choi (2015) used DEA to compare performance of production groups and individual products and detect potential difference among performance of product groups with ANOVA (Analysis of Variance) approach. They identified the source of inefficiency and validated their findings with Tobit regression analysis. Hwang (2016) proposed a cluster-based artificial contrasts approach to monitor non-homogeneously distributed data and showed that their approach was effective in reducing Type-II errors. Kumar et al. (2016) studied the data imbalance problem in a cloud-based manufacturing setting and developed a manufacturing platform enabled by MapReduce and cloud computing for fault diagnosis. They reported that it is essential to consider the effect of errors in data on product quality for cloud-based manufacturing. Shin et al. (2016) proposed a virtual machining model that captured process planning data for the generation of virtual records of machines' actions. Their model could be applied to assess performance of machines. Wang and Zhang (2016) developed a wafer lots' cycle time prediction approach and their experiments showed that their approach had a higher accuracy over linear regression and back-propagation network. Ahn and Lee (2017) proposed an artificial neural network approach to predict drilling burr formation. Chien, Liu, and Chuang (2017) offered a data-driven approach for excursion detection by analysing large volumes of data from semiconductor manufacturing processes.

4.1.3 *Product design*

A product design meeting customers' expectations can attract customers and boost sales. High-quality products tend to retain customers. The progress in information and communication technology and the product use data collected offer opportunities to further improve product design. Son et al. (2014) proposed a collaborative design environment offering a platform to share the information regarding the product designs to reduce lead time and identify design errors at early design stages. In an application of the modularity concept, Li and Xie (2015) developed an extraction algorithm to generate a design structure matrix from the CAD assembly information. They adopted four hierarchical clustering algorithms to generate modules. Arcidiacono et al. (2017) adopted Kriging models to optimise braking performance of freight trains, with the goal of minimising the adverse effects of in-train forces. Cui et al. (2017) proposed a neural network to optimise mass and radial deformation of high-pressure turbine discs.

4.1.4 *Environmental issues*

Sustainable manufacturing in an emerging research topic. Sun et al. (2014) applied a DEA approach to allocate emission permits to individual firms. They considered two schemes where the first one allowed one of the firms to allocate the emission permits and the other one employed a third party involvement. They found that the latter scheme maximised both the overall and individual efficiency. Wu, Zhu, et al. (2016) adopted DEA to allocate emission permits. Hsu, Lim, and Yang (2017) presented a fuel consumption prediction model based on drivers' driving behaviour and vehicle characteristics. Their prediction model could be used to aid vehicle design.

4.2. *Supply chain management*

Supply chains involve multiple parties, numerous objects and various types of processes. To keep track of the activities and events, products in a supply chain are frequently tagged with sensing devices for real-time tracing and tracking and further analysis. Recent applications include individuals as tracking objects, e.g. customers. The large volumes of data collected could be leveraged for better coordination among supply chain parties, more efficient communication, more agile response to changes and improved solutions for operational issues. In this section, the data-driven supply chain applications at the tactical and operational level are reviewed.

4.2.1 *Tactical level*

Empirical studies have demonstrated the benefits of embracing big data for supply chain management at a tactical level. Chae, Olson, and Sheu (2014) proposed a supply chain analytics framework including data management resources, IT-enabled planning resources and performance management resources. Using data collected from 537 manufacturing plants, the authors reported that data management resources are keys to business analytics initiatives. They also found that manufacturing firms which adopted more sophisticated planning technologies were more likely to benefit from data-driven practices. Wu, Chu, et al. (2016) considered the degree of satisfaction of a decision-making unit (DMU) towards a set of optimal weights for another DMU and developed a DEA cross-efficiency evaluation approach. They applied their methodology for technology selection. Using a sample of 4877 U.S. manufacturing companies, Hoberg, Badorf, and Lapp (2017) found that the fiscal calendar impact the inventory level, which the authors termed 'the inverse hockey stick effect'. Dotoli, Epicoco, and Falagario (2017) maximised the supply chain network efficiency under uncertainty. They proposed a three-phase approach to tackling the problem, where the first phase was to rank actors of each supply chain network, the second phase was to determine each stakeholders' required quantity, and the third was to limit the exchange of small quantities. Hofmann (2017) studied how big data can reduce the impact of bullwhip effect on a supply chain. The author found that among three big data characteristics – 'velocity', 'volume' and 'variety', – 'velocity' is the most significant for improving supply chain performance. Yao (2017) investigated the dynamic coordination and equilibrium between the service capacities from the supply and demand sides under a B2C online shopping setting. The author proposed an optimisation model for supply chain resources integration and developed an improved ant algorithm to provide a solution.

4.2.2 *Operational level*

4.2.2.1 *Logistics management.* Logistics management is one of the areas that could embrace big data, as various product data is collected for logistics purposes, e.g. real-time tracking. The data-enabled applications for logistics management are becoming more data-driven. Yayla et al. (2015) proposed an integrated approach of fuzzy techniques to third-party logistics service provider selection. Kengpol and Tuammee (2016) used a combination of techniques, including quantitative risk assessment, analytic hierarchy process and DEA to evaluate risks in a multimodal green logistics setting. Their methodology produces user-customised optimal green logistics routes. Pang and Chan (2017) used association rules to a storage location assignment problem. Pang and Gebka (2017) developed a forecasting model for the container throughput using SARIMA, seasonal Holt–Winters model and vector error correction model. Pjevcevic et al. (2017) employed DEA for the evaluation of the efficiency of container handling processes at a port container terminal. They established a simulation model to examine the performance measures and found that the fleet size of AGVs and efficiency evaluation of the operations were crucial for planning. Tsai and Huang (2017) proposed a neural network approach with factors including GDP, interest rates, amounts of import and export trades and quantities of containers and quay cranes to predict the flow of containers. van der Spoel, Amrit, and van Hillegersberg (2017) suggested a data-driven approach for cargo arrival time prediction at distribution centres. They studied 230 trucks, including the weather data to illustrate factors impacting the arrival times.

4.2.2.2 *Operations management.* Besides logistics, studies on the usage of data in managing processes in supply chains at an operational level have been published. Melnyk et al. (2014) addressed the issue that conventional simulation models provide steady-state results. However, one may need to react with transient responses, for example, in case of new product introduction and disruptions in supply chains. They integrated time series techniques and simulation modelling to determine transient responses. Dev et al. (2016) combined agent-based simulation and a decision tree learning algorithm to reconfigure the operational units of a mobile phone supply chain. Akcay and Corlu (2017) proposed a simulation replication algorithm that takes into account the uncertainty of the input when developing an inventory simulation model. Their methodology provided more accurate estimation of the service level compared with the conventional approaches such as the best-fit or empirical distribution as the demand distribution. Li and Wang (2017) applied the tracking and sensory technology for chilled food chain management. The data captured were leveraged to estimate the remaining shift-life of the perishable products.

4.3 *Customer research*

A better understanding of the customers can improve product design, enhance customer management and potentially increase the demand for a product, which has a direct linkage to the corporate success. Most customer research is data-driven, as the required knowledge for the research must be captured through customers' opinions and behaviours.

4.3.1 *Product design*

Meeting customer requirements (CRs) is critical to product design. Papinniemi, Hannola, and Maletz (2014) discussed challenges of requirements management relevant to production life cycle management in automotive industry. They proposed a framework integrating the two components for identifying interdependencies of product structures and requirements and examining the effects of life cycle changes. Jin et al. (2016) utilised large volumes of consumer opinion data to acquire information about the CRs. They proposed a Kalman filter method to predict the CR trend and a Bayesian framework to compare products. They also carried out a case study using consumer opinion data from Amazon and demonstrated how the data could be used for market-driven product design. Martí Bigorra and Isaksson (2017) considered the information regarding how a customer uses a product to enable more effective product design. Product usage data and customer needs were taken into account for design requirements. They adopted an analytical hierarchy process for risk assessment.

4.3.2 *Customer management*

There are different aspects of research in customer management. Abdolvand, Albadvi, and Aghdasi (2015) attempted to align performance management and customer management and proposed a model to maximise customer lifetime value. In their research, several techniques such as genetic k-means, analytical hierarchical process and DEA were adopted. Boongasame, Temdee, and Daneshgar (2017) studied the trust issues in buyer coalition schemes. They introduced the concepts of 'group signature' and 'authority'. Wu et al. (2017) considered consumers' preference to big data analytics in their model of competition between wearable device companies with and without adopting big data analytics. They found that when the consumers have different preferences, the market structure impacts the competition.

4.3.3 *Operational decisions*

The data on customer behaviour has been applied to optimise operational decisions. Tsai and Huang (2015) utilised data on customers' purchase and moving behaviours collected by RFID to optimise product space allocation decisions. Lee (2017) proposed a genetic algorithm to anticipate future shipping demands. The author proposed a cloud platform for processing large volumes of data arriving at various sources and adopted cluster-based association rule mining to identify purchase pattern and make predictions.

4.3.4 *Online platforms and social media*

Data collected from online platforms and social media is emerging. Such data can be acquired at a low cost, as it is often publicly accessible. While much non-trivial and valuable information can be potentially extracted from such data sources, the challenge in that the data is usually not well organised, unstructured and non-numerical (e.g. text, images). Researchers are beginning to investigate how such data could provide value in production research. Chan et al. (2017) discussed how social media data such as posts and comments could be leveraged for operations management research. They analysed information on social media and proposed a clustering approach to discover relationships among different factors. Chong et al. (2017) studied how factors such as online promotional marketing and reviews, discounts and delivery fee exemption influence the demands for electronic products sold at Amazon. They applied a neural network to discover that online reviews and promotions are important in demand prediction. Jiang et al. (2017) proposed a multi-class classification approach to identify useful quality-related reviews on social media. With a more effective identification of these quality-related reviews, it would be helpful for companies to enhance the quality of their products. Kim and Ahn (2017) integrated social network analysis and clustering algorithms into a collaborative filtering framework to identify the most influential people within an online social network. This information could be used to provide product recommendations to the customers.

4.4 ***Summary of production data research***

Figure 7 summarises the papers on production data research published from 2014 to 2017. Half of the papers were related to assessing performance of manufacturing systems and processes and quality monitoring. The use of data in supply chain management at an operational level has made the third category by the number of papers published in this period.

Discussion and future research directions are provided in the next section.

5. Discussion and future developments

The conventional production research methodologies tend to versed in *analytical models*. The analytical approaches require a prior understanding of the nature, characteristics and mechanisms of the problem. The data-sets are collected and processed so that models can be solved. The *data-driven* algorithms have become most widely used in production research. They do not require a prior understanding of the problem, but a large number of observations to produce a solution. These algorithms are particularly suitable for problems where relationships between variables have not been formally established, the required model inputs cannot be directly obtained from the available data-sets or the computational complexity of the analytical model is excessive.

To address the impact of the transformation from analytical to data-driven approaches, an in-depth discussion and future research directions are discussed next.

5.1 *Methodologies*

5.1.1 *Revisiting analytical modelling assumptions*

The development of analytical models requires assumptions. If the assumptions are not made properly, the findings and conclusions from the research may not reflect those in reality (Nasser and Tariq 2015). The wide variety and volumes of data captured have opened up a new opportunity for researchers to revisit these modelling assumptions. These assumptions may be validated or disproved. Unexplored relationships between variables can also be examined using big data. New modelling assumptions can be established or refined.

5.1.2 *Machine learning in production systems*

Data-driven approaches are gaining recognition in production research. The research on data-driven algorithms has been growing since 2000s. The applications of data-driven algorithms are also diverse, including product design (e.g. Agard and Kusiak 2004; Shao et al. 2006; Song and Kusiak 2009), improving manufacturing processes (e.g. Li, Chang, and Ni 2009; Yu and Matta 2016) and monitoring manufacturing activities (Jiang et al. 2014).

While the use of machine learning in production research dates to 1980s (e.g. Lu and Ham 1989; Shaw and Whinston 1989; Zaghloul et al. 1989), it has begun to gain acceptance in recent years. Machine learning algorithms can process symbolic data (e.g. texts, images and audio records), where analytical models are not likely to offer a viable approach. The applications of machine learning in production research are no longer restricted to the use for internal manufacturing operations, but also supply chain management (e.g. Dev et al. 2016), transportation (e.g. van der Spoel, Amrit, and van Hillegersberg 2017) and customer research (e.g. Jiang et al. 2017; Kim and Ahn 2017).

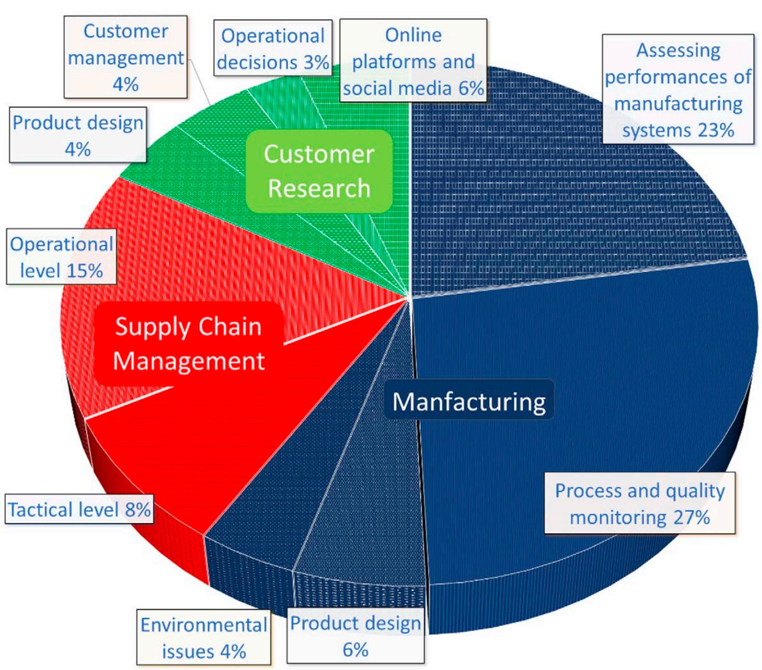

Figure 7. Production data research published in *IJPR* in the period of 2014–2017 by category.

5.1.3 *Data-intensive computational frameworks*

The developments in production data research have demonstrated that the methodologies adopted for data research are becoming more data-driven and data-intensive. Big data analytics has shown to be a key driver to offer business value (Ji-fan Ren et al. 2017). Conventional techniques are not suited to manage to handle high dimensionality data. While DEA remains one of the major techniques used in production data research, the applications of data mining are growing. More powerful computational solutions and platforms are required to enable new applications. For example, MapReduce and cloud computing frameworks have been adopted to facilitate computation (Kumar et al. 2016). There is a need to develop computing frameworks (Lee 2017; Zhong et al. 2017) in support of industrial implementations. The algorithms running on these computational platforms need to be more efficient, in particular for applications that require real-time solutions. It will be necessary to investigate how these computationally expensive problems can be scaled down. Feature selection, highly effective sampling techniques and problem decomposition with parallel computing can be potential solutions enabling a prompt response.

5.1.4 *Higher granularity simulation modelling of production systems*

The data collected from production systems can be used as input to process activity models, including virtual models (Lechevalier et al. 2017). Experiments with different inputs produce new data for further analysis. In the era of big data and IoT, high fidelity simulation models can be built. For example, customer behaviours can be captured by sensors such as RFID (Tsai and Huang 2015). Events, activities and characteristics that could not be captured in the past may now be considered. As an example, the impact of the retail facility layout, shelf allocation, pricing decisions and customer flow on the product demand may be simulated.

5.1.5 *Hybrid of analytical models and data-driven methodologies*

Integration of analytical and data-driven models could overcome the shortcomings of each of the two approaches. A partial understanding of the problem can serve as the foundation of the algorithms, whereas other hidden relationships can be explored by data-driven techniques. While there have been attempts to develop integrate multiple data-driven approaches in production (e.g. Jayaswal and Adil 2004; Abdolvand, Albadvi, and Aghdasi 2015), the research on hybrid of analytical and data-driven methodologies has shortcomings. The challenge of this line of research is to establish theoretical foundations and mechanisms to link the two types of techniques.

5.2 *Data*

5.2.1 *Integration of data sources from multiple parties in supply chains*

The volume of data captured in supply chains has been increasing (Addo-Tenkorang and Helo 2016; Zhong et al. 2016). Streaming makes the data dynamic. Data from different sources will make the processes and systems more coordinated, thereby resulting in higher productivity, efficiency and profitability. An integrated platform capturing data from parties in a supply chain and external information (customers' behaviour, social media and economic conditions) can respond to changes more proactively and disseminate immediate recommendations to relevant parties, for example, manufacturers to design a new product, logistics service providers to anticipate the shipping requests, and warehouses to reserve storage for the new product, in a more coordinated manner. However, such platform requires large data storage and powerful computing infrastructure. Parties in the supply chain may have different expectations for data sharing and therefore a trusted data platform becomes essential. Even if these parties can provide their own data, data aggregation remains difficult.

5.2.2 *Utilisation of external data sources*

The data-sets used in production research are usually collected and owned by the corporations. Limited by the nature of these internally captured data, the production research in the past was inclined to applications on the supply side. As discussed in Section 2, 'manufacturing', 'process' and 'control' were major focuses of production research in the past. Recently, there has been a growing trend that researchers utilise more external data sources for conducting the research. The data publicly available at online platforms and social media are an emerging source of new research directions. This has been addressed in the recent four papers on social media, Chan et al. (2017); Chong et al. (2017); Jiang et al. (2017); Kim and Ahn (2017). These external sources of information provide researchers with a more comprehensive picture of the factors impacting the product demand and may be utilised to design effective strategies for enhancing the demand (e.g. Kim and Ahn 2017).

5.2.3 *Utilisation of data-sets in heterogeneous formats*

The heterogeneous data formats, the natures of data – numerical, text, image and signal – and the problem of data from different sources not being synchronised (e.g. some collected each second, but some collected each hour) are challenging. Data standardisation in production is a must for future data research.

5.3 *Applications*

5.3.1 *Cyber–physical systems*

Smart manufacturing has been considered as one of the most encouraging future trends in production. This is indicated by the significant number of papers on this topic published in the recent years. The importance of smart manufacturing has been acknowledged by the industry.

Smart manufacturing aims to realise the concept of the integration of the physical assets in manufacturing systems and cyberspace to form cyber–physical systems (Kusiak 2018). It simultaneously utilises machines, sensors, data-driven models to make manufacturing processes 'smarter'. The adoption of cyber–physical systems for manufacturing has shown to be promising (Wang et al. 2015). The significance of smart manufacturing is largely dependent on the technological advances in manufacturing technology and the computational models. The enabling technologies for smart manufacturing include *Internet of Things (IoT)* and *cloud manufacturing*, which will be discussed later. To drive manufacturing systems smart, the essence is the development of mathematical models and computational algorithms that optimise integrated decisions across all physical assets, learning from historical manufacturing data and utilising real-time sensing information. Higher energy productivity can also be achieved (Edgar and Pistikopoulos 2017).

5.3.2 *Internet of things*

The integration of the physical manufacturing resources and cyberspace is developed based on the concept of IoT. IoT is the network of things (i.e. objects) – usually in the form of devices such as RFID tags, sensors, electronics and actuators – where the objects are connected and able to exchange information through information and communication infrastructures (Atzori, Iera, and Morabito 2010; Gubbi et al. 2013). IoT alone only provides information visibility, but does not offer the capability of automated decision-making (Bi, Da Xu, and Wang 2014). To enable smart manufacturing, the core intelligence needs to be captured by models and from data. Computationally efficient data-driven models become necessary in decision-making in a dynamic manufacturing environment. Furthermore, smart interconnection between the physical assets in manufacturing systems and cyberspace has to be established (Tao et al. 2017). Such smart interconnection covers all aspects of interactions between physical manufacturing resources and cyberspace: physical to physical (P2P), physical to cyber (P2C), cyber to cyber (C2C) and cyber to physical (C2P) (Tao, Cheng, and Qi 2017). The realisation of smart manufacturing requires effective and smooth communication and coordination between the physical manufacturing equipment (hardware) and the manufacturing intelligence (software) (Tao et al. 2014).

5.3.3 *Cloud manufacturing*

IoT is usually enabled by cloud computing platforms, or sometimes known as *platform as a service* (PaaS) (Tao et al. 2014). Cloud platforms enable enterprises and individual users to develop, test and run applications without the need of building and maintaining their own computing infrastructures. Cloud computing technologies have recently been adopted in manufacturing (Ren et al. 2017).

The adoption of cloud manufacturing can benefit the production processes in several aspects, including the ease of information exchange among multiple parties in the production process (Talhi et al. 2017), sharing of manufacturing resources (Tao et al. 2017) and smart distributed scheduling (He and Xu 2015; Zhang et al. 2017). From an economic aspect, cloud manufacturing also offers companies flexibility of scaling up and down their production costs because of the 'pay-as-you-go' business models (Xu 2012). The keys of enabling cloud computing include a comprehensive and compute-intensive infrastructure, fast and flexible configuration on the computing platforms, capability of exchanging high volumes of information in real time, scalability in platform architecture, standardised data formats and trusted and secured data sharing.

5.3.4 *Personalised production research*

Data related to individual customers has been more frequently utilised in production research. Such information captured at an individual level (e.g. from social media or customers' purchasing records) provides an opportunity to understand the customers better. Product design, product recommendations and marketing strategies can be more personalised. New data-driven models based on text mining and clustering algorithms for personalised production research is required to achieve the goals.

There may be a need to rethink how mass customisation shall be implemented to anticipate the impacts of this higher degree of personalisation on manufacturing processes. *Additive manufacturing* technologies, have become mature and increasingly ubiquitous in manufacturing systems. The technologies can enable rapid and on-demand manufacturing. Customers may also submit requests for their own product designs through cloud platforms. How these personalised production requests are handled (e.g. which components shall be produced by 3D printing), in combination with mass production plans, in an automated and coordinated manner requires further investigation.

5.3.5 *Predictive maintenance*

Predictive maintenance has drawn more attention in recent years. With cyber–physical systems, data are captured through various sensing devices supporting prognostics and health management (Xia and Xi 2017). Machine learning has been applied for predictions of errors in different types of manufacturing activities (e.g. Li, Wang, and Wang 2017; Nikula et al. 2017; Ramos et al. 2017). Data regarding non-manufacturing activities can also be utilised to predict other types of failure in the production process, such as disruptions in logistics activities (Gürbüz et al. 2017). The formats of data captured often heterogeneous: numerical (e.g. weight of the product, temperature of the machine and pollutant levels), image (e.g. appearance of the product), audio (e.g. sounds during the manufacturing process). The analytics performed on this integrated sources of information shall improve the accuracy for estimating the likelihood of machine breakdowns. However, the challenges are the data aggregation of heterogeneous data-sets.

5.3.6 *Sustainability of production systems*

The recent developments of production research enabled by data have tried to address the environmental issues resulting from production processes (e.g. Sun et al. 2014; Wu, Zhu, et al. 2016; Hsu, Lim, and Yang 2017). Sustainability issues have become more and more of a concern to manufacturing companies. As an example, to address the pressing environmental issues in China, the Chinese government has recently established environmental laws (Forbes 2017). Big data is an emerging facilitator to guide manufacturing processes to be more environmentally friendly. For instance, data about manufacturing and maintenance processes can be analysed in real time for deriving strategies for cleaner manufacturing and supporting better product life cycle management (Zhang et al. 2017). Data-driven approaches can also be applied for green supplier selection (Shabanpour, Yousefi, and Saen 2017). Recent research also suggests that big data analytics has the impacts on world-class sustainable manufacturing (Dubey et al. 2016). While research has been conducted for sustainable production, there is still room to implement the concepts and assessing their impacts in real life (Chun and Bidanda 2013).

6. Conclusion

Applications of data and data analysis methodologies in production have been studied for over five decades. The review of papers published by *IJPR* relevant to production data indicated that the research was sporadic before 1990, became regular but not major around 1990s, and intensified around 2000 due to advances in data capturing and processing technologies. The research has shifted from analytical models to data-driven approaches. The major applications were in analysis and improvement of processes, systems, products and management. Analysis has been a primary goal of the research, but analytics (utilising advanced mathematics, statistics and computer science) has gained recognition in recent years.

In terms of solution methodologies, data envelopment analysis (DEA) was the most widely adopted methodology in production data research, but its popularity has declined recently. It was demonstrated that data mining algorithms gained popularity due to the growing volumes and wider variety of production data. Big data is an emerging topic in production research. New types and sources of data, such as social media, have provided researchers with opportunities to explore new areas in production research.

The production data research was classified into three categories – manufacturing, supply chain management and customer research, where research on manufacturing has been the major part. Data-driven and machine learning algorithms were applied in many of these applications.

This paper discussed challenges and future opportunities in data research, including the data itself, methodologies and applications. In addition to the traditional challenges in production research (such as model and algorithm developments), data processing and computational platforms and integration of data of heterogeneous formats from multiple sources are new concerns. Nevertheless, exciting applications have realised, e.g. cyber–physical systems enabled by Internet of Things (IoT) and cloud manufacturing, personalised production research, predictive maintenance and more sustainable production systems.

Acknowledgement

This research is partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 14202115).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Yong-Hong Kuo http://orcid.org/0000-0002-6170-324X

References

Abbasimehr, H., M. Setak, and J. Soroor. 2013. “A Framework for Identification of High-value Customers by including Social Network Based Variables for Churn Prediction Using Neuro-fuzzy Techniques.” *International Journal of Production Research* 51 (4): 1279–1294.

Abdolvand, N., A. Albadvi, and M. Aghdasi. 2015. “Performance Management Using a Value-based Customer-centered Model.” *International Journal of Production Research* 53 (18): 5472–5483.

Aberg, U. 1967. “An Approach to the Systematic Arrangement of Work Measurement Data.” *International Journal of Production Research* 6 (4): 257–268.

Addo-Tenkorang, R., and P. T. Helo. 2016. “Big Data Applications in Operations/Supply-chain Management: A Literature Review.” *Computers & Industrial Engineering* 101: 528–543.

Agard, B., and A. Kusiak. 2004. “Data-mining-based Methodology for the Design of Product Families.” *International Journal of Production Research* 42 (15): 2955–2969.

Ahn, Y., and S. H. Lee. 2017. “Classification and Prediction of Burr Formation in Micro Drilling of Ductile Metals.” *International Journal of Production Research* 55 (17): 4833–4846.

Akcay, A., and C. G. Corlu. 2017. “Simulation of Inventory Systems with Unknown Input Models: A Data-driven Approach.” *International Journal of Production Research* 55 (19): 5826–5840.

Arcidiacono, G., R. Berni, L. Cantone, and P. Placidoli. 2017. “Kriging Models for Payload Distribution Optimisation of Freight Trains.” *International Journal of Production Research* 55 (17): 4878–4890.

Atzori, L., A. Iera, and G. Morabito. 2010. “The Internet of Things: A Survey.” *Computer Networks* 54 (15): 2787–2805.

Azadeh, A., T. Nazari, and H. Charkhand. 2015. “Optimisation of Facility Layout Design Problem with Safety and Environmental Factors by Stochastic DEA and Simulation Approach.” *International Journal of Production Research* 53 (11): 3370–3389.

Azadeh, A., M. Sheikhalishahi, M. Firoozi, and S. M. Khalili. 2013. “An Integrated Multi-criteria Taguchi Computer Simulation-DEA Approach for Optimum Maintenance Policy and Planning by Incorporating Learning Effects.” *International Journal of Production Research* 51 (18): 5374–5385.

Bao, L., K. Wang, and R. Jin. 2014. “A Hierarchical Model for Characterising Spatial Wafer Variations.” *International Journal of Production Research* 52 (6): 1827–1842.

Bayraktar, E., A. Gunasekaran, S. L. Koh, E. Tatoglu, M. Demirbag, and S. Zaim. 2010. “An Efficiency Comparison of Supply Chain Management and Information Systems Practices: A Study of Turkish and Bulgarian Small-and Medium-Sized Enterprises in Food Products and Beverages.” *International Journal of Production Research* 48 (2): 425–451.

Ben-Arieh, D., and E. Triantaphyllou. 1992. “Quantifying Data for Group Technology with Weighted Fuzzy Features.” *International Journal of Production Research* 30 (6): 1285–1299.

Bi, Z., L. Da Xu, and C. Wang. 2014. “Internet of Things for Enterprise Systems of Modern Manufacturing.” *IEEE Transactions on Industrial Informatics* 10 (2): 1537–1546.

Bischak, D. P., and E. A. Silver. 2004. “Estimating the Out-of-control Rate from Control Chart Data in the Presence of Multiple Causes and Process Improvement.” *International Journal of Production Research* 42 (24): 5217–5233.

Boongasame, L., P. Temdee, and F. Daneshgar. 2017. “A Group Signature Based Buyer Coalition Scheme with Trustable Third Party.” *International Journal of Production Research* 55 (17): 5050–5061.

Boyer, S. A. 2009. “SCADA: Supervisory Control and Data Acquisition.” *International Society of Automation* [published online].

Braha, D., Y. Elovici, and M. Last. 2007. “Theory of Actionable Data Mining with Application to Semiconductor Manufacturing Control.” *International Journal of Production Research* 45 (13): 3059–3084.

Brintrup, A., D. Ranasinghe, and D. McFarlane. 2010. “RFID Opportunity Analysis for Leaner Manufacturing.” *International Journal of Production Research* 48 (9): 2745–2764.

Broomhead, P., and M. Edkins. 1986. “Generating NC Data at the Machine Tool for the Manufacture of Free-form Surfaces.” *International Journal of Production Research* 24 (1): 1–14.

Buddhakulsomsiri, J., Y. Siradeghyan, A. Zakarian, and X. Li. 2006. “Association Rule-generation Algorithm for Mining Automotive Warranty Data.” *International Journal of Production Research* 44 (14): 2749–2770.

Buyya, R., C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic. 2009. “Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility.” *Future Generation Computer Systems* 25 (6): 599–616.

Caramia, M., and G. Felici. 2006. “Mining Relevant Information on the Web: A Clique-based Approach.” *International Journal of Production Research* 44 (14): 2771–2787.

Chae, B., D. Olson, and C. Sheu. 2014. “The Impact of Supply Chain Analytics on Operational Performance: A Resource-based View.” *International Journal of Production Research* 52 (16): 4695–4710.

Chakravorty, S. S., and D. N. Hales. 2017. “Sustainability of Process Improvements: An Application of the Experiential Learning Model (ELM).” *International Journal of Production Research* 55 (17): 4931–4947.

Chan, H. K., E. Lacka, R. W. Yee, and M. K. Lim. 2017. “The Role of Social Media Data in Operations and Production Management.” *International Journal of Production Research* 55 (17): 5027–5036.

Chandrasekharan, M. P., and R. Rajagopalan. 1989. “GROUPABIL1TY: An Analysis of the Properties of Binary Data Matrices for Group Technology.” *International Journal of Production Research* 27 (6): 1035–1052.

Chang, S. Y., and T. H. Chen. 2008. “Performance Ranking of Asian Lead Frame Firms: A Slack-Based Method in Data Envelopment Analysis.” *International Journal of Production Research* 46 (14): 3875–3885.

Chang, D. S., Y. C. Kuo, and T. Y. Chen. 2008. “Productivity Measurement of the Manufacturing Process for Outsourcing Decisions: The Case of a Taiwanese Printed Circuit Board Manufacturer.” *International Journal of Production Research* 46 (24): 6981–6995.

Chen, M. C., C. C. Hsu, B. Malhotra, and M. K. Tiwari. 2016. “An Efficient ICA-DW-SVDD Fault Detection and Diagnosis Method for Non-Gaussian Processes.” *International Journal of Production Research* 54 (17): 5208–5218.

Chen, M. C., C. T. Su, C. C. Hsu, and Y. W. Liu. 2005. “Data Transformation in SPC for Semiconductor Machinery Control: A Case of Monitoring Particles.” *International Journal of Production Research* 43 (13): 2759–2773.

Chen, K. S., K. J. Wang, and T. C. Chang. 2017. “A Novel Approach to Deriving the Lower Confidence Limit of Indices C Pu, C Pl, and C Pk in Assessing Process Capability.” *International Journal of Production Research*: 1–19.

Chiang, Y. M., and F. L. Chen. 1999. “Sculptured Surface Reconstruction from CMM Measurement Data by a Software Iterative Approach.” *International Journal of Production Research* 37 (8): 1679–1695.

Chien, C. F., S. C. Hsu, and Y. J. Chen. 2013. “A System for Online Detection and Classification of Wafer Bin Map Defect Patterns for Manufacturing Intelligence.” *International Journal of Production Research* 51 (8): 2324–2338.

Chien, C. F., C. W. Liu, and S. C. Chuang. 2017. “Analysing Semiconductor Manufacturing Big Data for Root Cause Detection of Excursion for Yield Enhancement.” *International Journal of Production Research* 55 (17): 5095–5107.

Chinnam, R. B., and P. Baruah. 2009. “Autonomous Diagnostics and Prognostics in Machining Processes through Competitive Learning-driven HMM-based Clustering.” *International Journal of Production Research* 47 (23): 6739–6758.

Chong, A. Y. L., E. Ch’ng, M. J. Liu, and B. Li. 2017. “Predicting Consumer Product Demands via Big Data: The Roles of Online Promotional Marketing and Online Reviews.” *International Journal of Production Research* 55 (17): 5142–5156.

Chou, C. Y., C. H. Chen, and H. R. Liu. 2001. “Economic Design of X Charts for Non-normally Correlated Data.” *International Journal of Production Research* 39 (9): 1931–1941.

Chun, Y., and B. Bidanda. 2013. “Sustainable Manufacturing and the Role of the International Journal of Production Research.” *International Journal of Production Research* 51 (23–24): 7448–7455.

Colosimo, B. M., and M. Pacella. 2010. “A Comparison Study of Control Charts for Statistical Monitoring of Functional Data.” *International Journal of Production Research* 48 (6): 1575–1601.

Cui, D., G. Feng, P. Zhou, and Y. Zhang. 2017. “Parametric Design-based Multi-objective Optimisation for High-pressure Turbine Disc.” *International Journal of Production Research* 55 (17): 4847–4861.

Da Cunha, C., B. Agard, and A. Kusiak. 2006. “Data Mining for Improvement of Product Quality.” *International Journal of Production Research* 44 (18–19): 4027–4041.

Da Cunha, C., B. Agard, and A. Kusiak. 2010. “Selection of Modules for Mass Customisation.” *International Journal of Production Research* 48 (5): 1439–1454.

Das, D., S. Zhou, Y. Chen, and J. Horst. 2016. “Statistical Monitoring of Over-dispersed Multivariate Count Data Using Approximate Likelihood Ratio Tests.” *International Journal of Production Research* 54 (21): 6579–6593.

Davenport, T. H., and J. Dyché (2013). “Big Data in Big Companies.” *International Institute for Analytics* [published online].

Dengiz, O., A. E. Smith, and I. Nettleship. 2006. “Two-stage Data Mining for Flaw Identification in Ceramics Manufacture.” *International Journal of Production Research* 44 (14): 2839–2851.

Dev, N. K., R. Shankar, A. Gunasekaran, and L. S. Thakur. 2016. “A Hybrid Adaptive Decision System for Supply Chain Reconfiguration.” *International Journal of Production Research* 54 (23): 7100–7114.

Ding, D., F. Tsung, and J. Li. 2016. “Directional Control Schemes for Processes with Mixed-type Data.” *International Journal of Production Research* 54 (6): 1594–1609.

Dotoli, M., and M. Falagario. 2012. “A Hierarchical Model for Optimal Supplier Selection in Multiple Sourcing Contexts.” *International Journal of Production Research* 50 (11): 2953–2967.

Dotoli, M., N. Epicoco, and M. Falagario. 2017. “A Fuzzy Technique for Supply Chain Network Design with Quantity Discounts.” *International Journal of Production Research* 55 (7): 1862–1884.

Dubey, R., A. Gunasekaran, S. J. Childe, S. F. Wamba, and T. Papadopoulos. 2016. “The Impact of Big Data on World-class Sustainable Manufacturing.” *The International Journal of Advanced Manufacturing Technology* 84 (1–4): 631–645.

Edgar, T. F., and E. N. Pistikopoulos. 2017. “Smart Manufacturing and Energy Systems.” *Computers & Chemical Engineering* [published online].

Evans, L., N. Lohse, K. H. Tan, P. Webb, and M. Summers. 2012. “Justification for the Selection of Manufacturing Technologies: A Fuzzy-decision-tree-based Approach.” *International Journal of Production Research* 50 (23): 6945–6962.

Feng, J. X., and A. Kusiak. 2006. “Data Mining Applications in Engineering Design, Manufacturing and Logistics.” *International Journal of Production Research* 44 (14): 2689–2694.

Forbes. 2017. *China Shuts down Tens of Thousands of Factories in Widespread Pollution Crackdown*. Accessed 13 December 2017. https://www.forbes.com/sites/trevornace/2017/10/24/china-shuts-down-tens-of-thousands-of-factories-in-widespread-pollution-crackdown/

Foroughi, A. A. 2012. “A Modified Common Weight Model for Maximum Discrimination in Technology Selection.” *International Journal of Production Research* 50 (14): 3841–3846.

Forza, C., and F. Salvador. 2008. “Application Support to Product Variety Management.” *International Journal of Production Research* 46 (3): 817–836.

Franklin Liu, F. H., and Y. Cheng Liu. 2008. “Product Line Performance Assessment on Order Fulfilment Cycle Time: A Case of Microelectronic Communication Company.” *International Journal of Production Research* 46 (16): 4431–4443.

Godinho Filho, M., A. G. Marchesini, J. Riezebos, N. Vandaele, and G. M. D. Ganga. 2017. “The Extent of Knowledge of Quick Response Manufacturing Principles: An Exploratory Transnational Study.” *International Journal of Production Research* 55 (17): 4891–4911.

Grasso, M., B. M. Colosimo, and M. Pacella. 2014. “Profile Monitoring via Sensor Fusion: The Use of PCA Methods for Multi-Channel Data.” *International Journal of Production Research* 52 (20): 6110–6135.

Grasso, M., B. M. Colosimo, and F. Tsung. 2017. “A Phase I Multi-modelling Approach for Profile Monitoring of Signal Data.” *International Journal of Production Research* 55 (15): 4354–4377.

Gubbi, J., R. Buyya, S. Marusic, and M. Palaniswami. 2013. “Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions.” *Future Generation Computer Systems* 29 (7): 1645–1660.

Guh, R. S., Y. R. Shiue, and T. Y. Tseng. 2011. “The Study of Real Time Scheduling by an Intelligent Multi-controller Approach.” *International Journal of Production Research* 49 (10): 2977–2997.

Gunaratne, N. G. T., M. A. Abdollahian, S. Huda, and J. Yearwood. 2017. “Exponentially Weighted Control Charts to Monitor Multivariate Process Variability for High Dimensions.” *International Journal of Production Research* 55 (17): 4948–4962.

Gupta, T., and H. I. Seifoddini. 1990. “Production Data Based Similarity Coefficient for Machine-component Grouping Decisions in the Design of a Cellular Manufacturing System.” *International Journal of Production Research* 28 (7): 1247–1269.

Gürbüz, F., İ. Eski, B. Denizhan, and C. Dağlı. 2017. “Prediction of Damage Parameters of a 3PL Company via Data Mining and Neural Networks.” *Journal of Intelligent Manufacturing* 1: 1–13.

Hasani, A., S. H. Zegordi, and E. Nikbakhsh. 2012. “Robust Closed-loop Supply Chain Network Design for Perishable Goods in Agile Manufacturing Under Uncertainty.” *International Journal of Production Research* 50 (16): 4649–4669.

He, Q. P., and J. Wang. 2017. “Statistical Process Monitoring as a Big Data Analytics Tool for Smart Manufacturing.” *Journal of Process Control* [published online].

He, W., and L. Xu. 2015. “A State-of-the-Art Survey of Cloud Manufacturing.” *International Journal of Computer Integrated Manufacturing* 28 (3): 239–250.

He, Z., L. Zuo, M. Zhang, and F. M. Megahed. 2016. “An Image-based Multivariate Generalized Likelihood Ratio Control Chart for Detecting and Diagnosing Multiple Faults in Manufactured Products.” *International Journal of Production Research* 54 (6): 1771–1784.

Hoberg, K., F. Badorf, and L. Lapp. 2017. “The Inverse Hockey Stick Effect: An Empirical Investigation of the Fiscal Calendar’s Impact on Firm Inventories.” *International Journal of Production Research* 55 (16): 4601–4624.

Hoffer, J. A. 2011. *Modern Database Management*. 10th ed. Pearson Education India.

Hofmann, E. 2017. “Big Data and Supply Chain Decisions: The Impact of Volume, Variety and Velocity Properties on the Bullwhip Effect.” *International Journal of Production Research* 55 (17): 5108–5126.

Hou, J. L., and S. T. Yang. 2006. "Technology-mining Model concerning Operation Characteristics of Technology and Service Providers." *International Journal of Production Research* 44 (16): 3345–3365.

Hsu, C. Y., S. S. Lim, and C. S. Yang. 2017. "Data Mining for Enhanced Driving Effectiveness: An Eco-driving Behaviour Analysis Model for Better Driving Decisions." *International Journal of Production Research* 55 (23): 7096–7109.

Huang, C. C., T. L. Tseng, H. F. Chuang, and H. F. Liang. 2006. "Rough-set-based Approach to Manufacturing Process Document Retrieval." *International Journal of Production Research* 44 (14): 2889–2911.

Hwang, W. Y. 2016. "Cluster-based Artificial Contrasts for Inhomogeneously Distributed Data with an Indicator Variable." *International Journal of Production Research* 54 (17): 5045–5055.

Hwarng, H. B. 1995. "Multilayer Perceptions for Detecting Cyclic Data on Control Charts." *International Journal of Production Research* 33 (11): 3101–3117.

Jain, S., G. Shao, and S. J. Shin. 2017. "Manufacturing Data Analytics Using a Virtual Factory Representation." *International Journal of Production Research* 55 (18): 5450–5464.

Jayaswal, S., and G. K. Adil. 2004. "Efficient Algorithm for Cell Formation with Sequence Data, Machine Replications and Alternative Process Routings." *International Journal of Production Research* 42 (12): 2419–2433.

Jeong, M. K., J. C. Lu, and N. Wang. 2006. "Wavelet-based SPC Procedure for Complicated Functional Data." *International Journal of Production Research* 44 (4): 729–744.

Jiang, C., Y. Liu, Y. Ding, K. Liang, and R. Duan. 2017. "Capturing Helpful Reviews from Social Media for Product Quality Improvement: A Multi-class Classification Approach." *International Journal of Production Research* 55 (12): 3528–3541.

Jiang, Q., X. Yan, Z. Lv, and M. Guo. 2014. "Independent Component Analysis-based Non-Gaussian Process Monitoring with Preselecting Optimal Components and Support Vector Data Description." *International Journal of Production Research* 52 (11): 3273–3286.

Ji-fan Ren, S., S. Fosso Wamba, S. Akter, R. Dubey, and S. J. Childe. 2017. "Modelling Quality Dynamics, Business Value and Firm Performance in a Big Data Analytics Environment." *International Journal of Production Research* 55 (17): 5011–5026.

Jin, Y., and Y. Ishino. 2006. "DAKA: Design Activity Knowledge Acquisition through Data-mining." *International Journal of Production Research* 44 (14): 2813–2837.

Jin, J., Y. Liu, P. Ji, and H. Liu. 2016. "Understanding Big Consumer Opinion Data for Market-driven Product Design." *International Journal of Production Research* 54 (10): 3019–3041.

Jin, X., B. W. Wah, X. Cheng, and Y. Wang. 2015. "Significance and Challenges of Big Data Research." *Big Data Research* 2 (2): 59–64.

Johansson, B., A. Skoogh, J. Andersson, K. Ahlberg, and L. Hanson. 2015. "Power-level Sampling of Metal Cutting Machines for Data Representation in Discrete Event Simulation." *International Journal of Production Research* 53 (23): 7060–7070.

Joo, S. J., and H. C. P. Choi. 2015. "Assessing Comparative Production Efficiencies for Product Line Management." *International Journal of Production Research* 53 (17): 5241–5250.

Kacar, N. B., and R. Uzsoy. 2014. "A Comparison of Multiple Linear Regression Approaches for Fitting Clearing Functions to Empirical Data." *International Journal of Production Research* 52 (11): 3164–3184.

Kanagawa, A., F. Tamaki, and H. Ohta. 1993. "Control Charts for Process Average and Variability Based on Linguistic Data." *International Journal of Production Research* 31 (4): 913–922.

Kang, H. S., J. Y. Lee, S. Choi, H. Kim, J. H. Park, J. Y. Son, B. H. Kim, and S. Do Noh. 2016. "Smart Manufacturing: Past Research, Present Findings, and Future Directions." *International Journal of Precision Engineering and Manufacturing-Green Technology* 3 (1): 111–128.

Kang, C. W., M. B. Ramzan, B. Sarkar, and M. Imran. 2018. "Effect of Inspection Performance in Smart Manufacturing System Based on Human Quality Control System." *The International Journal of Advanced Manufacturing Technology* 94 (9–12): 4351–4364.

Kendrick, C. 1961. "Estimation of Service Requirements for Production Purposes." *International Journal of Production Research* 1 (1): 1–12.

Kengpol, A., and S. Tuammee. 2016. "The Development of a Decision Support Framework for a Quantitative Risk Assessment in Multimodal Green Logistics: An Empirical Study." *International Journal of Production Research* 54 (4): 1020–1038.

Kim, K. J., and H. Ahn. 2017. "Parametric Empirical Bayes Estimation of Individual Time-pressure Reactivity." *International Journal of Production Research* 644: 1–12.

Kim, Y. S., and W. J. Kolarik. 1992. "Real-time Conditional Reliability Prediction from On-line Tool Performance Data." *International Journal of Production Research* 30 (8): 1831–1844.

Kim, S., and D. A. Nembhard. 2013. "Rule Mining for Scheduling Cross Training with a Heterogeneous Workforce." *International Journal of Production Research* 51 (8): 2281–2300.

Kitchin, R. 2014. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* [published online].

Kumar, A., R. Shankar, A. Choudhary, and L. S. Thakur. 2016. "A Big Data MapReduce Framework for Fault Diagnosis in Cloud-based Manufacturing." *International Journal of Production Research* 54 (23): 7060–7073.

Kuo, R. J., and Y. J. Lin. 2012. "Supplier Selection Using Analytic Network Process and Data Envelopment Analysis." *International Journal of Production Research* 50 (11): 2852–2863.

Kusiak, A. 1987. "The Generalized Group Technology Concept." *International Journal of Production Research* 25 (4): 561–569.

Kusiak, A. 2006. “Data Mining: Manufacturing and Service Applications.” *International Journal of Production Research* 44 (18–19): 4175–4191.

Kusiak, A. 2017. “Smart Manufacturing Must Embrace Big Data.” *Nature* 544 (7648): 23–25.

Kusiak, A. 2018. “Smart Manufacturing.” *International Journal of Production Research* http://www.tandfonline.com/doi/full/10.1080/00207543.2017.1351644.

Kwak, D. S., K. J. Kim, and M. S. Lee. 2010. “Multistage PRIM: Patient Rule Induction Method for Optimisation of a Multistage Manufacturing Process.” *International Journal of Production Research* 48 (12): 3461–3473.

Lechevalier, D., S. J. Shin, S. Rachuri, S. Foufou, Y. T. Lee, and A. Bouras. 2017. “Simulating a Virtual Machining Model in an Agent-based Model for Advanced Analytics.” *Journal of Intelligent Manufacturing* [published online].

Lee, C. K. H. 2017. “A GA-based Optimisation Model for Big Data Analytics Supporting Anticipatory Shipping in Retail 4.0.” *International Journal of Production Research* 55 (2): 593–605.

Lee, A. C., D. P. Chen, and C. L. Lin. 1990. “A CAD/CAM System from 3D Coordinate Measuring Data.” *International Journal of Production Research* 28 (12): 2353–2371.

Lee, J. Y., M. Swink, and T. Pandejpong. 2017. “Team Diversity and Manufacturing Process Innovation Performance: The Moderating Role of Technology Maturity.” *International Journal of Production Research* 55 (17): 4912–4930.

Li, L. 2009. “Bottleneck Detection of Complex Manufacturing Systems Using a Data-driven Method.” *International Journal of Production Research* 47 (24): 6929–6940.

Li, L., Q. Chang, and J. Ni. 2009. “Data Driven Bottleneck Detection of Manufacturing Systems.” *International Journal of Production Research* 47 (18): 5019–5036.

Li, D. C., W. C. Chen, C. J. Chang, C. C. Chen, and I. H. Wen. 2015. “Practical Information Diffusion Techniques to Accelerate New Product Pilot Runs.” *International Journal of Production Research* 53 (17): 5310–5319.

Li, D. C., and W. L. Dai. 2009. “Determining the Optimal Collaborative Benchmarks in a Supply Chain.” *International Journal of Production Research* 47 (16): 4457–4471.

Li, F., G. C. Runger, and E. Tuv. 2006. “Supervised Learning for Change-point Detection.” *International Journal of Production Research* 44 (14): 2853–2868.

Li, J. R., and Q. H. Wang. 2010. “A Rough Set Based Data Mining Approach for House of Quality Analysis.” *International Journal of Production Research* 48 (7): 2095–2107.

Li, D., and X. Wang. 2017. “Dynamic Supply Chain Decisions Based on Networked Sensor Data: An Application in the Chilled Food Retail Chain.” *International Journal of Production Research* 55 (17): 5127–5141.

Li, Z., Y. Wang, and K. Wang. 2017. “A Data-driven Method Based on Deep Belief Networks for Backlash Error Prediction in Machining Centers.” *Journal of Intelligent Manufacturing* [published online].

Li, D. C., I. H. Wen, and W. C. Chen. 2016. “A Novel Data Transformation Model for Small Data-set Learning.” *International Journal of Production Research* 54 (24): 7453–7463.

Li, B. M., and S. Q. Xie. 2015. “Module Partition for 3D CAD Assembly Models: A Hierarchical Clustering Method Based on Component Dependencies.” *International Journal of Production Research* 53 (17): 5224–5240.

Liao, M. Y. 2015. “Assessing Process Incapability When Collecting Data from Multiple Batches.” *International Journal of Production Research* 53 (7): 2041–2054.

Liu, J. Y., M. Xie, T. N. Goh, and L. Y. Chan. 2007. “A Study of EWMA Chart with Transformed Exponential Data.” *International Journal of Production Research* 45 (3): 743–763.

Liu, C., K. Yasuda, Y. Yin, and K. Tanaka. 2009. “Uncertain Association Rule Mining Algorithm for the Cell Formation Problem in Cellular Manufacturing Systems.” *International Journal of Production Research* 47 (3): 667–685.

Lu, S. C., and I. Ham. 1989. “Machine Learning Techniques for Group Technology Applications.” *CIRP Annals* 38 (1): 455–459.

Lu, W. M., and S. W. Hung. 2010. “Assessing the Performance of a Vertically Disintegrated Chain by the DEA Approach – A Case Study of Taiwanese Semiconductor Firms.” *International Journal of Production Research* 48 (4): 1155–1170.

Lu, W. M., W. K. Wang, and H. L. Lee. 2013. “The Relationship between Corporate Social Responsibility and Corporate Performance: Evidence from the US Semiconductor Industry.” *International Journal of Production Research* 51 (19): 5683–5695.

Manikas, A., L. Boyd, Q. Pang, and J. Guan. 2017. “An Analysis of Research Methods in IJPR since Inception.” *International Journal of Production Research* [published online].

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity.* Seattle, WA: McKinsey Global Institute.

Martí Bigorra, A., and O. Isaksson. 2017. “Combining Customer Needs and the Customer’s Way of Using the Product to Set Customer-focused Targets in the House of Quality.” *International Journal of Production Research* 55 (8): 2320–2335.

Melnyk, S. A., C. W. Zobel, J. R. Macdonald, and S. E. Griffis. 2014. “Making Sense of Transient Responses in Simulation Studies.” *International Journal of Production Research* 52 (3): 617–632.

Metan, G., I. Sabuncuoglu, and H. Pierreval. 2010. “Real Time Selection of Scheduling Rules and Knowledge Extraction via Dynamically Controlled Data Mining.” *International Journal of Production Research* 48 (23): 6909–6938.

Mittal, S., Khan, M. A., Romero ,D., and Wuest ,T. 2017. Smart Manufacturing: Characteristics, Technologies and Enabling Factors. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* [published online].

Mousavi, A., and H. R. A. Siervo. 2017. "Automatic Translation of Plant Data into Management Performance Metrics: A Case for Real-time and Predictive Production Control." *International Journal of Production Research* 55 (17): 4862–4877.

Nair, G. J. 1999. "Accord: A Bicriterion Algorithmfor Cell Formation Using Ordinal and Ratio-level Data." *International Journal of Production Research* 37 (3): 539–556.

Nair, G. J., and T. T. Narendran. 1998. "CASE: A Clustering Algorithm for Cell Formation with Sequence Data." *International Journal of Production Research* 36 (1): 157–180.

Nasser, T., and R. S. Tariq. 2015. "Big Data Challenges." *Journal of Computer Engineering & Information Technology* 4 (3) [published online].

Neaga, E. I., and J. A. Harding. 2005. "An Enterprise Modeling and Integration Framework Based on Knowledge Discovery and Data Mining." *International Journal of Production Research* 43 (6): 1089–1108.

Ngai, E. W. 2010. "RFID Technology and Applications in Production and Supply Chain Management." *International Journal of Production Research* 48 (9): 2481–2483.

Ngai, E. W., A. Gunasekaran, S. F. Wamba, S. Akter, and R. Dubey. 2017. "Big Data Analytics in Electronic Markets." *Electronic Markets* 27 (3): 243–245.

Nikula, R. P., K. Karioja, K. Leiviskä, and E. Juuso. 2017. "Prediction of Mechanical Stress in Roller Leveler Based on Vibration Measurements and Steel Strip Properties." *Journal of Intelligent Manufacturing* [published online].

Noh, C. H., S. L. Lee, D. H. Kim, C. W. Chung, and S. H. Kim. 2010. "An Effective and Efficient Defect Inspection System for TFT-LCD Polarised Films Using Adaptive Thresholds and Shape-based Image Analyses." *International Journal of Production Research* 48 (17): 5115–5135.

O'Keefe, R. M., and J. Haddock. 1991. "Data-driven Generic Simulators for Flexible Manufacturing Systems." *International Journal of Production Research* 29 (9): 1795–1810.

Omar, R. S. M., U. Venkatadri, C. Diallo, and S. Mrishih. 2017. "A Data-driven Approach to Multi-Product Production Network Planning." *International Journal of Production Research* 55 (23): 7110–7134.

Pang, K. W., and H. L. Chan. 2017. "Data Mining-based Algorithm for Storage Location Assignment in a Randomised Warehouse." *International Journal of Production Research* 55 (14): 4035–4052.

Pang, G., and B. Gebka. 2017. "Forecasting Container Throughput Using Aggregate or Terminal-Specific Data? The Case of Tanjung Priok Port, Indonesia." *International Journal of Production Research* 55 (9): 2454–2469.

Papinniemi, J., L. Hannola, and M. Maletz. 2014. "Challenges in Integrating Requirements Management with PLM." *International Journal of Production Research* 52 (15): 4412–4423.

Parmar, D., T. Wu, T. Callarman, J. Fowler, and P. Wolfe. 2010. "A Clustering Algorithm for Supplier Base Management." *International Journal of Production Research* 48 (13): 3803–3821.

Parthiban, P., H. A. Zubar, and P. Katakar. 2013. "Vendor Selection Problem: A Multi-criteria Approach Based on Strategic Decisions." *International Journal of Production Research* 51 (5): 1535–1548.

Pjevcevic, D., M. Nikolic, N. Vidic, and K. Vukadinovic. 2017. "Data Envelopment Analysis of AGV Fleet Sizing at a Port Container Terminal." *International Journal of Production Research* 55 (14): 4021–4034.

Qian, Z., W. Jiang, and K. L. Tsui. 2006. "Churn Detection via Customer Profile Modelling." *International Journal of Production Research* 44 (14): 2913–2933.

Quariguasi Frota Neto, J., G. Walther, J. A. E. E. Bloemhof, J. A. E. E. Van Nunen, and T. Spengler. 2010. "From Closed-loop to Sustainable Supply Chains: The WEEE Case." *International Journal of Production Research* 48 (15): 4463–4481.

Raheja, D., J. Llinas, R. Nagi, and C. Romanowski. 2006. "Data Fusion/Data Mining-based Architecture for Condition-based Maintenance." *International Journal of Production Research* 44 (14): 2869–2887.

Ramos, A. R., J. M. B. de Lázaro, A. Prieto-Moreno, A. J. da Silva Neto, and O. Llanes-Santiago. 2017. "An Approach to Robust Fault Diagnosis in Mechanical Systems Using Computational Intelligence." *Journal of Intelligent Manufacturing* [published online].

Ren, L., L. Zhang, L. Wang, F. Tao, and X. Chai. 2017. "Cloud Manufacturing: Key Characteristics and Applications." *International Journal of Computer Integrated Manufacturing* 30 (6): 501–515.

Riegler, M., N. André, M. Gronalt, and T. M. Young. 2015. "Real-time Dynamic Simulation of Continuous Bulk Material Flow to Improve the Statistical Modelling of Final Product Strength Properties." *International Journal of Production Research* 53 (21): 6629–6636.

Sarkis, J., and J. Dijkshoorn. 2007. "Relationships between Solid Waste Management Performance and Environmental Practice Adoption in Welsh Small and Medium-sized Enterprises (SMEs)." *International Journal of Production Research* 45 (21): 4989–5015.

Sevkli, M., S. C. Lenny Koh, S. Zaim, M. Demirbag, and E. Tatoglu. 2007. "An Application of Data Envelopment Analytic Hierarchy Process for Supplier Selection: A Case Study of BEKO in Turkey." *International Journal of Production Research* 45 (9): 1973–2003.

Shabanpour, H., S. Yousefi, and R. F. Saen. 2017. "Forecasting Efficiency of Green Suppliers by Dynamic Data Envelopment Analysis and Artificial Neural Networks." *Journal of Cleaner Production* 142: 1098–1107.

Shafaei, R., and P. Brunn. 1999a. "Workshop Scheduling Using Practical (Inaccurate) Data Part 1: The Performance of Heuristic Scheduling Rules in a Dynamic Job Shop Environment Using a Rolling Time Horizon Approach." *International Journal of Production Research* 37 (17): 3913–3925.

Shafaei, R., and P. Brunn. 1999b. "Workshop Scheduling Using Practical (Inaccurate) Data Part 2: An Investigation of the Robustness of Scheduling Rules in a Dynamic and Stochastic Environment." *International Journal of Production Research* 37 (18): 4105–4117.

Shao, X. Y., Z. H. Wang, P. G. Li, and C. X. J. Feng. 2006. "Integrating Data Mining and Rough Set for Customer Group-based Discovery of Product Configuration Rules." *International Journal of Production Research* 44 (14): 2789–2811.

Shaw, M. J., and A. B. Whinston. 1989. "An Artificial Intelligence Approach to the Scheduling of Flexible Manufacturing Systems." *IIE Transactions* 21 (2): 170–183.

Shin, S. J., J. Woo, D. B. Kim, S. Kumaraguru, and S. Rachuri. 2016. "Developing a Virtual Machining Model to Generate MTConnect Machine-Monitoring Data from STEP-NC." *International Journal of Production Research* 54 (15): 4487–4505.

Shiue, Y. R. 2009. "Data-mining-based Dynamic Dispatching Rule Selection Mechanism for Shop Floor Control Systems Using a Support Vector Machine Approach." *International Journal of Production Research* 47 (13): 3669–3690.

Shore, H. 1998. "A New Approach to Analysing Non-normal Quality Data with Application to Process Capability Analysis." *International Journal of Production Research* 36 (7): 1917–1933.

Skinner, K. R., D. C. Montgomery, and G. C. Runger. 2003. "Process Monitoring for Multiple Count Data Using Generalized Linear Model-based Control Charts." *International Journal of Production Research* 41 (6): 1167–1180.

Son, S., S. Na, K. Kim, and S. Lee. 2014. "Collaborative Design Environment between ECAD and MCAD Engineers in High-tech Products Development." *International Journal of Production Research* 52 (20): 6161–6174.

Song, Z., and A. Kusiak. 2009. "Optimising Product Configurations with a Data-mining Approach." *International Journal of Production Research* 47 (7): 1733–1751.

van der Spoel, S., C. Amrit, and J. van Hillegersberg. 2017. "Predictive Analytics for Truck Arrival Time Estimation: A Field Study at a European Distribution Centre." *International Journal of Production Research* 55 (17): 5062–5078.

Sun, J., J. Wu, L. Liang, R. Y. Zhong, and G. Q. Huang. 2014. "Allocation of Emission Permits Using DEA: Centralised and Individual Points of View." *International Journal of Production Research* 52 (2): 419–435.

Sun, S. 2002. "Assessing Computer Numerical Control Machines Using Data Envelopment Analysis." *International Journal of Production Research* 40 (9): 2011–2039.

Talhi, A., V. Fortineau, J. C. Huet, and S. Lamouri. 2017. "Ontology for Cloud Manufacturing Based Product Lifecycle Management." *Journal of Intelligent Manufacturing* [published online].

Talluri, S., F. Huq, and W. E. Pinney. 1997. "Application of Data Envelopment Analysis for Cell Performance Evaluation and Process Improvement in Cellular Manufacturing." *International Journal of Production Research* 35 (8): 2157–2170.

Talluri, S., R. Narasimhan, and S. Viswanathan. 2007. "Information Technologies for Procurement Decisions: A Decision Support System for Multi-attribute E-reverse Auctions." *International Journal of Production Research* 45 (11): 2615–2628.

Tan, K. H., G. Ji, C. P. Lim, and M. L. Tseng. 2017. "Using Big Data to Make Better Decisions in the Digital Economy." *International Journal of Production Research* 55 (17): 4998–5000.

Tao, F., Y. Cheng, L. Da Xu, L. Zhang, and B. H. Li. 2014. "CCIoT-CMfg: Cloud Computing and Internet of Things-based Cloud Manufacturing Service System." *IEEE Transactions on Industrial Informatics* 10 (2): 1435–1442.

Tao, F., J. Cheng, and Q. Qi. 2017. "IIHub: An Industrial Internet-of-things Hub towards Smart Manufacturing Based on Cyber-physical System." *IEEE Transactions on Industrial Informatics* [published online].

Tao, F., Y. Cheng, L. Zhang, and A. Y. Nee. 2017. "Advanced Manufacturing Systems: Socialization Characteristics and Trends." *Journal of Intelligent Manufacturing* 28 (5): 1079–1094.

Towill, D. R. 1973. "A Direct Method for the Determination of Learning Curve Parameters from Historical Data." *International Journal of Production Research* 11 (1): 97–101.

Towill, D. R. 1977. "Exponential Smoothing of Learning Curve Data." *International Journal of Production Research* 15 (1): 1–15.

Tsai, C. Y., and S. H. Huang. 2015. "A Data Mining Approach to Optimise Shelf Space Allocation in Consideration of Customer Purchase and Moving Behaviours." *International Journal of Production Research* 53 (3): 850–866.

Tsai, F. M., and L. J. Huang. 2017. "Using Artificial Neural Networks to Predict Container Flows Between the Major Ports of Asia." *International Journal of Production Research* 55 (17): 5001–5010.

Tseng, T. L., C. C. Huang, F. Jiang, and J. C. Ho. 2006. "Applying a Hybrid Data-mining Approach to Prediction Problems: A Case of Preferred Suppliers Prediction." *International Journal of Production Research* 44 (14): 2935–2954.

Vainio, F., M. Maier, T. Knuutila, E. Alhoniemi, M. Johnsson, and O. S. Nevalainen. 2010. "Estimating Printed Circuit Board Assembly times Using Neural Networks." *International Journal of Production Research* 48 (8): 2201–2218.

Wan, H. D., and F. Frank Chen. 2008. "A Leanness Measure of Manufacturing Systems for Quantifying Impacts of Lean Initiatives." *International Journal of Production Research* 46 (23): 6567–6584.

Wang, F. K. 2005. "A Simple Data Transformation of Auto-Correlated Data for SPC." *International Journal of Production Research* 43 (5): 981–989.

Wang, W. K., Y. C. Chan, W. M. Lu, and H. Chang. 2015. "The Impacts of Asset Impairments on Performance in the Taiwan Listed Electronics Industry." *International Journal of Production Research* 53 (8): 2410–2426.

Wang, Y. M., and K. S. Chin. 2009. "A New Approach for the Selection of Advanced Manufacturing Technologies: DEA with Double Frontiers." *International Journal of Production Research* 47 (23): 6663–6679.

Wang, Y. M., K. S. Chin, and J. P. F. Leung. 2009. "A Note on the Application of the Data Envelopment Analytic Hierarchy Process for Supplier Selection." *International Journal of Production Research* 47 (11): 3121–3138.

Wang, C. H., and Y. W. Chien. 2016. "Combining Balanced Scorecard with Data Envelopment Analysis to Conduct Performance Diagnosis for Taiwanese LED Manufacturers." *International Journal of Production Research* 54 (17): 5169–5181.

Wang, R. T., C. T. B. Ho, and K. Oh. 2010. "Measuring Production and Marketing Efficiency Using Grey Relation Analysis and Data Envelopment Analysis." *International Journal of Production Research* 48 (1): 183–199.

Wang, L., M. Törngren, and M. Onori. 2015. "Current Status and Advancement of Cyber-physical Systems in Manufacturing." *Journal of Manufacturing Systems* 37: 517–527.

Wang, C. H., S. J. Wang, and W. D. Lee. 2006. "Automatic Identification of Spatial Defect Patterns for Semiconductor Manufacturing." *International Journal of Production Research* 44 (23): 5169–5185.

Wang, J., and J. Zhang. 2016. "Big Data Analytics for Forecasting Cycle Time in Semiconductor Wafer Fabrication System." *International Journal of Production Research* 54 (23): 7231–7244.

Wei, G., and J. Wang. 2011. "Value-at-Risk and Data Envelopment Analysis: Comments on Wu and Olson (2010)." *International Journal of Production Research* 49 (23): 7189–7193.

Wong, J. T., C. T. Su, and H. T. Hsieh. 2007. "A Framework for Determining MIMO Process Parameters by a Neuro-DMandACO Approach." *International Journal of Production Research* 45 (15): 3505–3520.

Wortham, A. W. 1972. "The Use of Exponentially Smoothed Data in Continuous Process Control." *International Journal of Production Research* 10 (4): 393–400.

Wu, T., and J. Blackhurst. 2009. "Supplier Evaluation and Selection: An Augmented DEA Approach." *International Journal of Production Research* 47 (16): 4593–4608.

Wu, D., and D. L. Olson. 2008. "A Comparison of Stochastic Dominance and Stochastic DEA for Vendor Evaluation." *International Journal of Production Research* 46 (8): 2313–2327.

Wu, D. D., and D. Olson. 2010. "Enterprise Risk Management: A DEA VaR Approach in Vendor Selection." *International Journal of Production Research* 48 (16): 4919–4932.

Wu, D. D., and D. Olson. 2011. "A Note to 'Enterprise Risk Management: A DEA VaR Approach in Vendor Selection': A Response to Wei and Wang and Model Extension." *International Journal of Production Research* 49 (23): 7195–7203.

Wu, J., J. Chu, Q. Zhu, P. Yin, and L. Liang. 2016. "DEA Cross-efficiency Evaluation Based on Satisfaction Degree: An Application to Technology Selection." *International Journal of Production Research* 54 (20): 5990–6007.

Wu, J., H. Li, Z. Lin, and K. Y. Goh. 2017. "How Big Data and Analytics Reshape the Wearable Device Market – The Context of E-health." *International Journal of Production Research* 55 (17): 5168–5182.

Wu, J., Q. Zhu, J. Chu, Q. An, and L. Liang. 2016. "A DEA-Based Approach for Allocation of Emission Reduction Tasks." *International Journal of Production Research* 54 (18): 5618–5633.

Xia, T., and L. Xi. 2017. "Manufacturing Paradigm-oriented PHM Methodologies for Cyber-physical Systems." *Journal of Intelligent Manufacturing* [published online].

Xu, X. 2012. "From Cloud Computing to Cloud Manufacturing." *Robotics and Computer-Integrated Manufacturing* 28 (1): 75–86.

Yao, J. M. 2017. "Supply Chain Resources Integration Optimisation in B2C Online Shopping." *International Journal of Production Research* 55 (17): 5079–5094.

Yayla, A. Y., A. Oztekin, A. T. Gumus, and A. Gunasekaran. 2015. "A Hybrid Data Analytic Methodology for 3PL Transportation Provider Evaluation Using Fuzzy Multi-criteria Decision Making." *International Journal of Production Research* 53 (20): 6097–6113.

Ye, N. 2017. "A Reverse Engineering Algorithm for Mining a Causal System Model from System Data." *International Journal of Production Research* 55 (3): 828–844.

Yu, F. J., C. Y. Low, and S. S. Cheng. 2003. "Design for an SPRT Control Scheme Based on Linguistic Data." *International Journal of Production Research* 41 (6): 1299–1309.

Yu, C., and A. Matta. 2016. "A Statistical Framework of Data-driven Bottleneck Identification in Manufacturing Systems." *International Journal of Production Research* 54 (21): 6317–6332.

Zaghloul, M. E., D. Khera, L. W. Linholm, and C. P. Reeve. 1989. "A Machine-learning Classification Approach for IC Manufacturing Control Based on Test Structure Measurements." *IEEE Transactions on Semiconductor Manufacturing* 2 (2): 47–53.

Zhang, J., G. Ding, Y. Zou, S. Qin, and J. Fu. 2017. "Review of Job Shop Scheduling Research and its New Perspectives under Industry 4.0." *Journal of Intelligent Manufacturing* [published online].

Zhang, X., C. K. M. Lee, and S. Chen. 2012. "Supplier Evaluation and Selection: A Hybrid Model Based on DEAHP and ABC." *International Journal of Production Research* 50 (7): 1877–1889.

Zhang, Y., S. Ren, Y. Liu, and S. Si. 2017. "A Big Data Analytics Architecture for Cleaner Manufacturing and Maintenance Processes of Complex Products." *Journal of Cleaner Production* 142: 626–641.

Zhong, R. Y., S. T. Newman, G. Q. Huang, and S. Lan. 2016. "Big Data for Supply Chain Management in the Service and Manufacturing Sectors: Challenges, Opportunities, and Future Perspectives." *Computers & Industrial Engineering* 101: 572–591.

Zhong, R. Y., C. Xu, C. Chen, and G. Q. Huang. 2017. "Big Data Analytics for Physical Internet-based Intelligent Manufacturing Shop Floors." *International Journal of Production Research* 55 (9): 2610–2621.