# A Review of Supply Chain Data Mining Publications

David L. Olson [*]

*College of Business Administration, University of Nebraska – Lincoln, Lincoln, NE 68588-0491*

Published first: 30-06-2020

**Abstract** – The use of data mining in supply chains is growing, and covers almost all aspects of supply chain management. A framework of supply chain analytics is used to classify data mining publications reported in supply chain management academic literature. Scholarly articles were identified using SCOPUS and EBSCO Business search engines. Articles were classified by supply chain function. Additional papers reflecting technology, to include RFID use and text analysis were separately reviewed. The paper concludes with discussion of potential research issues and outlook for future development.

**Keywords:** Data mining; Supply Chain Management; Review paper

## 1. Introduction

Supply Chain Analytics (SCA) refers to the use of data and quantitative tools and techniques to improve operational performance (Handfield 2006, Davis-Sramek et al. 2010). Analytics in supply chain management (SCM) is not necessarily a new idea (Davenport and O'Dwyer 2011), since various quantitative techniques and modeling methods have long been used in manufacturing firms. The recent surge of interest in SCA is accompanied by new challenges and opportunities in both business and information technology (IT) environments. These challenges include issues arising from managing large amounts of data (e.g., data availability, data quality) and dealing with environmental uncertainties (Handfield and Nichols 2004, Liberatore and Luo 2010, Lavalle et al. 2011).

IT is an important corollary to SCA. IT-based innovations have generated and captured more data while also changing the nature of businesses (Kohli and Grover 2008). The opportunity to gain competitive advantage may thus arise from how firms manage data (Forslund and Jonsson 2007). Another major challenge for businesses is the increasing uncertainty in both demand (e.g., consumer market) and supply sides of their businesses. Dealing with demand and supply uncertainty by means of proper supply chain planning has been a major theme in many recent SCM studies (Oliva and Watson 2011, Demirkan and Delen 2012).

This paper classifies supply chain applications of data mining, big data, and other forms of knowledge management across all functions of supply chains. The diversity of applications is widespread, and is expected to grow with the growth of technology. The paper also looks at additional reported applications more general than specific supply chain functions, highlighting technology in the form of RFID application, and special data mining tools to include text mining.

Data mining is a key component of business analytics, although there is much more to business analytics than the data mining component. We first define terms and their relationships. We then provide our view of supply chain management, and how business analytics (and specifically data mining) have been reported in academic literature. We also discuss technology development and categorize research addressing its role in supporting

---

[*] Corresponding author. Email address: dolson3@unl.edu.

supply chain management. The paper concludes with a review of how we see trends in development, and problems and opportunities in the further use of business analytics (focusing on data mining) in supply chains.

## 2. Framework for Analysis

This paper uses definition of terms, noting that many others have differing specifics although what is presented is compatible with most, at least at the higher levels. Knowledge management is a high level concept, which Chae et al. (2014) define as consisting of access to data and its management along with analytics. INFORMS (Robinson et al. 2010) gives the concise definition: "Analytics facilitates realization of business objectives through reporting of data to analyze trends, creating predictive models for forecasting and optimizing business processes for enhanced performance." Sathi (2012) gives two common sources of data grouped under the banner of Big Data - data within the corporation that, thanks to automation and access, is increasingly shared, and data from outside the organization - some available publicly free of cost, some based on paid subscription, and the rest available selectively for specific business partners or customers. Big data provides the opportunity to monitor many more things at a much more detailed level.

Knowledge Management is used here as an umbrella term, including subsets of applications useful to the supply chain firm (process management, etc.), the information systems needed to provide information for supply chain management, data sources, and the broad field of analytics. Table 1 displays these relationships.

Table 1. Definition of terms

| Knowledge management | | Elaboration |
|---|---|---|
| Performance management resources | How things are done (tacit knowledge, BPR) | Process control |
| | | Six Sigma |
| Information systems | Database, reports, decision support | Cloud computing |
| Data sources | ERP & related systems | RFID |
| | External sources | Government publications |
| | Big data | Social media |
| Management Science (Analytics) | Descriptive analysis | |
| | Data mining | Classification |
| | | Prediction |
| | | Clustering |
| | | Link analysis |
| | | Text mining |
| | Operations Research | Mathematical programming |
| | | Simulation |
| | | Stochastic modeling |

SCA has been proposed as a promising approach to better manage data, utilize information technology (IT) resources, and prepare for effective supply chain planning. This new generation of analytic tools can develop a firm's IT and data management capabilities to enhance planning and improve operational performance (Mithas et al. 2011). It is suggested that firms can use SCA from data acquisition (e.g., RFID) and repository (e.g., ERP) technologies to improve supply chain planning through IT-enabled planning and scheduling systems (O'Dwyer and Renner 2011).

Chae et al. (2014) took a theoretical perspective of SCA as a combination of three sets of data and IT-enabled supply chain management resources, to include data management resources (DMR), IT-based supply chain planning resources (IPR), and performance management resources (PMR). Manufacturing firms acquire and use various IT and organizational resources in these three aspects of SCA. DMR includes IT-related resources (e.g., RFID, ERP) and analytical capabilities (e.g., mathematical optimization techniques) for data acquisition and management. IPR indicates the level of sophistication of IT resources used for supply chain planning. PMR refers to a set of data-oriented resources focusing on improving manufacturing process performance, such as statistical process control and Six Sigma, among others.

The three subsequent subsections deal with papers covering the broader context of knowledge management, big data, and specific supply chain functions. Knowledge Management and Big Data sections are shorter because the focus is on specific supply chain functions supported by data mining.

## 2.1.  Knowledge Management Papers

Overview papers related to data analytics and/or knowledge management within supply chains discuss conceptual views of how knowledge management supports supply chains. Collins et al. (2010) gave an historical review of supply chain development along with a list of supply chain measures in the context of overall firm strategy. They found that by capturing data and mining it, firms could better identify environmental changes and to develop strategies for successful coping. Chae et al. (2014) analyzed the effectiveness of SCA in leading to greater manufacturing effectiveness. Banerjee et al. (2013) reported that state of the art of data analysis in India with respect to supply chain management as well as other areas. Choudhary et al. (2011) gave a specific system to support knowledge acquisition, learning, and updating within supply chains. Qin and Zhao (2012) discussed how cloud computing can aid supply chain information management.

## 2.2.  Big Data Papers

Big data has become endemic in our society. New regulatory requirements on food safety and tracking in the European Union and in the US are driving development of data platforms and advanced analytics throughout the supply chain (Dobbs et al. 2015). It is important in supply chain management because of a number of sources of data. RFID and bar-coding generate a great deal of data. Chae (2015) examined the role of social media platforms in generating big data useful within supply chains. Ivanov and Sokolov (2012) provided an inter-disciplinary modeling framework for model-based decision support within collaborative supply chains. Sensors and transaction logs within organizations, as well as Internet linkages both from fellow supply chain participants as well as from other Internet sources have led to an explosive growth in big data, which can be valuable to track materiel by location, and can be useful for supply chain planning.

Tien (2012) argued for big data as an enabler for mass customization of goods and/or services, leading to simultaneous real-time management of supply and demand chains by supply chain participants. This was contended to provide overwhelming economic advantage to industrialized countries where consumers can influence producers through big data feedback. Schlegel (2014/2015) discussed big data and predictive analytics as strategic tools to mitigate risk within supply chains, utilizing the Internet-of-things and data mining to better forecast demand. Schoenherr and Speier-Pero (2015) reviewed academic programs that have evolved to train university students in supply chain data science (predictive analytics and big data).

## 2.3.  Supply Chain Functions

Supply chain operations include many functions. A list of typical activities, not intended to be comprehensive, but representative:
Supply Chain Functions
- o Forecasting
- o Product development
  - ▪ Profiling         customer relationship management
- o Supplier selection
- o Inventory management        Order batching
- o Logistics
- o Quality management

This bullet list will be used as a framework for sorting out published articles related to data mining within supply chains. Fraud detection is a function that might appear in a specific supply chain, but not necessarily in all. Risk management is a comprehensive dimension to all aspects of supply chain management. We encountered papers in these additional areas as well as the core supply chain functions.

## 3. Supply Chain Function Papers

The initial source of papers were obtained from search engines EBSCO Business Premier and SCOPUS, using search terms "supply chain management" and "data mining." SCOPUS provided about twice as many returns, with many overlaps. These sources were supplemented by additional searches on specific topics. It is not possible to obtain a count by category with any meaning, as most articles cover multiple aspects of both knowledge management as well as of application to different supply chain functions. The review will commence with the seven categories outlines in section 2.3.

### 3.1. Forecasting

Forecasting is a well-studied area, long-important in data mining research. Accurately forecasting inventory replenishment requirements is critical for successful supply chain management (Stefanovic, 2015), and data mining can be used to do that better. Smith (2011) gave four examples of supply chain management involving forecasting. Thomassey (2010) gave a widespread example of the need for accurate forecasting in supply chain environments (in his case, textile apparel). The reported application included complex features such as fuzzy logic and neural networks, and were evaluated for mitigation of the bullwhip effect by simulation. Shukla et al. (2012) dealt with the common demand feature of seasonality, specifically investigating unintended cyclic variability from endogenous disturbances. Ability to identify these early would have obvious benefits to supply chain participants. This paper again applied advanced techniques, in this case Fourier and discrete wavelet transforms along with an autoregressive model. The Fourier transform was most effective in the steel supply network examined in that study. Tirkel (2013) applied forecasting to semiconductor manufacturing flow time analysis, utilizing data mining processes to include business understanding, data processing, and classification to forecast flow time, enabling better short-term planning and efficiency.

Related to forecasting is prediction. A prediction may be a statement of an expected outcome, while forecast is a term often used for generation of a range of possible outcomes. Thomassey (2010) applied fuzzy logic and neural networks to improve forecasting in a clothing industry application to reduce bullwhip effect. Stefanovic (2014) used the term "predictive analytics" in his paper encouraging supply chains to be proactive and to apply data mining for prediction. His business intelligence model included process modeling, performance measurement, data mining models, and web portal technologies, seeking more intelligent, predictive, and responsive supply chains.

### 3.2. Product Development

Product development is a key step in any production operation, and is thus important in many supply chains. Wang et al. (2004) presented a case-based reasoning model for product development. The paper deals with knowledge management rather than data mining, but serves as a demonstration of the difference. The tacit design knowledge of the participants were to be enhanced by the system pulling up past processes and potentially including support of tacit information. Data mining, conversely, deals more with quantitative analysis.

In the area of data mining, Tien (2006) classified methods used to manage supply and demand chains. They suggest that simultaneous management of supply and demand can foster mass customization in real-time. Song and Kusiak (2009) applied mass customization through data mining to obtain rules from historical sales data to reach compromise between customer preference for greater product diversity with production's time-honored emphasis on product line simplification.

In the review of papers, it was apparent that profiling of customers often blends into the same topic as product development. This seems appropriate in a more service-oriented environment than existed maybe 60 years ago during the heyday of the rust belt (where the focus was on "making things"). Three papers were originally categorized as product development, but clearly have customer relationship focus. Liao et al. (2008a) demonstrated use of the a priori algorithm of association rules to map products to customer demands with the intent of suggesting new products for specific customers. Liao et al. (2008b) added k-means clustering to the a priori algorithm to mine customer knowledge. Liao and Wen (2009) reported knowledge extraction from a priori association rules in new product development of bicycles in Taiwan.

Data mining has long been applied to customer relationship management. Review of these papers confirms its continued use, but also show that data mining can be used for more operational level activities in product development. Min (2006) demonstrated data mining to examine grocery customer behavior and gave a classic example of loyal customer profiling. Liao et al. (2009) used their *a priori*/clustering combination to market segmentation in the Taiwanese life insurance market. Pillutla et al. (2014) analytically mined surveys obtained from Chinese firms in Suzhou on adoption of e-procurement, enabling them to identify critical drivers for system adoption. Srivhok and Intrapairot (2014) used classification analysis to study tourist demand in Thailand.

### 3.3. Supplier Selection

Selecting suppliers seems to be the most commonly analyzed supply chain problem, at least with respect to data mining. Piramuthu (2005) provided an early framework using machine learning for automated supply chain configuration, which could be applied in real time over the Internet. Ni et al. (2007) gave a methodology for supplier selection including quality function deployment, focusing on robust analysis and decision support. Lin et al. (2009) used association rules to consider thousands of potential suppliers while considering a number of factors. Parmar et al. (2010) applied cluster analysis to filter suppliers down to more manageable groups. Zhao and Yu (2011) used case-based reasoning and data mining (for attribute weight calculation in entropy) and clustering along with back-propagation neural networks in the Chinese petroleum industry. Chen et al. (2012) reported an integrated model combining k-means clustering and decision trees to overcome limitations of statistical techniques demonstrated on a manufacturing application. Jain et al. (2014) applied data mining to identify hidden relationships among supplier pre-qualification data with overall supplier rating based on prior work to improve supplier selection.

Some research has considered specific multiple criteria aspects (as well as data mining) in supplier selection. Ha and Krishnan (2008) gave a method incorporating multiple techniques through a combined supplier score considering multiple dimensions of performance. Hong and Ha (2008) gave a partner assessment system to monitor supplier performance over time, utilizing machine learning and multiple criteria analysis, and demonstrated their system in the agricultural industry. Xu and Lin (2009) used analytic hierarchy process along with association rule mining, using a supplier selection scenario for demonstration.

There also are a number of papers utilizing fuzzy and related approaches. Jain et al. (2007) applied fuzzy association rule mining for supplier evaluation. Li et al. (2008) proposed grey-based rough sets along with data mining and rough set theory, arguing that it had relative advantages under conditions of high uncertainty. Jain et al. (2009) analyzed diverse modeling methods to include fuzzy logic, agent technology, and petri nets for dynamic supply chain partner configuration. Zou et al. (2011) used rough set theory to derived decision rules for distributor selection in China. Faezy Razi (2014) used a combination of grey relational analysis and an artificial bee colony heuristic algorithm to select suppliers under conditions of multiple criteria.

Supplier selection is probably the most popular research topic concerning analytic modeling in supply chain management. Thus it is expected that there are quite a few data mining papers. These papers appear to apply the gamut of data mining approaches, as well as knowledge management ideas such as case-based reasoning. These spill over into multiple criteria and fuzzy papers, of which there are many non-data mining papers in supplier selection research.

### 3.4. Inventory Management

Inventory management also has received a great deal of study in the supply chain field. Thus we are not surprised to find a number of data mining papers applied to inventory management. Song and Dinwoodie (2008) claimed data mining in their application of ANOVA analysis and stochastic dynamic programming to vendor management inventory reduction of lead-time. Tsou (2013) developed an autoregressive time series data mining technique to detect market demand change. Ji et al. (2013) proposed a system gathering data from Internet-of-things technology, fed into data mining algorithms to control supply chain inventory.

In the field of more traditional data mining tools (classification), Chen and Wu (2005) applied association rules along with zero-one programming to demand side management of electrical utilities. Delen et al. (2011) combined data mining models with operations research tools and geographic information systems for a US Department of Defense funded projects related to blood supply management.

Shao et al. (2010) used system dynamics, multi-agent modeling, and complex adaptive systems for forecasting supplier demand data. Peng et al. (2014) also applied system dynamics, in this case for disaster relief support to better manage inventory stocking.

Inventory management has long been a staple research topic in operations management, and continues to be so with the shift to focus on supply chain management. Forecasting has obvious utility, but classification modeling also is present. Simulation is highly appropriate due to the presence of variance in inventory.

### 3.5.  Logistics

There were fewer papers that directly seemed related to logistics, although this is clearly an important field in supply chain management. Bhattacharya et al. (2014) reported an application of support vector machines along with mixed integer programming for intermodal transportation support of traffic flow. Renaud-Gentié et al. (2014) used multidimensional data analysis for vineyard transportation network analysis in the Loire Valley of France. Soeanu et al. (2015) used decision trees to model transportation decisions manage transportation risk in the field of humanitarian aid.

Mathematical programming appears a bit more in logistics papers. Risk aspects are clearly appropriate, especially in the context of supply chains.

### 3.6.  Quality Management

The emphasis on quality function deployment in operations management makes it a rich field for application. Zhou et al. (2008) reported effective uses of on-line reputation systems, a form of text analysis, in supporting market management. Ho et al. (2009) gave a distributed process mining system that they contended would lead to shorter delivery time, greater flexibility, and higher customer satisfaction. Li et al. (2010) applied a data mining framework for early warning and proactive control of a food supply chain network. Ting et al. (2014) used association rules in their decision support system to monitor food quality, using Hong Kong red wine products as a demonstration application.

Ghadge et al. (2012) utilized text mining to analyze supply chain risk literature, claiming an unbiased and holistic picture of advances within supply chain risk management.

Text analysis clearly has a role in analyzing customer quality feedback. But analysis of other aspects related to supply chain quality also are evident.

### 3.7.  Risk Analysis

Supply chains usually involve trading off the control of vertical integration for the opportunities of combining the core competencies of a number of independent entities, often through outsourcing. Inherently, this loss of control involves more risk but lower cost. Management of supply chain risk is crucial. Application of supply chain risk management appeared in a number of articles already classified. Le et al. (2013) specifically applied association rules to retail supply chain to remove sensitive knowledge from released databases, a common issue in contemporary collaborative supply chain databases containing sensitive information. Another risk aspect is fraud detection, widely used in the insurance industry. Kraus and Valverde (2014) discussed data warehouse design supporting forensic analysis to detect fraud in inventory management and warranty claims, but argued that the methodology was generalizable to the supply chain management process.

Risk is a component of all aspects of supply chain management, so many risk-related articles appear in prior categories. The data variability inherent in risk makes simulation or stochastic modeling attractive.

## 4. Papers Dealing with Related Tools

A number of papers have been published focusing on the technologies enabling application of data mining (or more generally knowledge management) to supply chains. This section demonstrates one source of technology support to supply chains in the form of radio-frequency identification (RFID), and then looks at technology support in the form of data mining methodologies, with specific categorization of text mining.

## 4.1. RFID Papers

RFID offers great potential within supply chains, enabling critically important tracking capabilities that can greatly improve the efficiency of inventory management, as well as the management of other resources. Key challenges to the implementation of data mining using RFID data are that different identification numbers across participating supply chain organizations, and aggregation of data often limit traceability. Furthermore, because of the multitude of pallets that might be tracked, the sheer dimensionality of tracking is a challenge.

Thiesse et al. (2009) gave fundamental concepts and applications of a system based on Electronic Product Code that facilitates data exchange within supply chains, and thus enables data mining application. Gonzalez et al. (2010) addressed the massive data sets generated by RFID within supply chains, and provided a movement graph model to compactly represent RFID data sets. Lee et al. (2010) proposed a knowledge discovery system to support customer relationship management based on point-of-sales and historical data as well as RFID data. Their system queries temporal data interactively and recursively, followed by association rule modeling for market basket analysis. Yűksel and Yűksel (2011) introduced a middleware to link RFID technology automatically to supply chain systems. Masciari (2012) presented a system monitoring RFID data for objects tracked in supply chains. Brandau and Tolujevs (2013) suggested data mining methods to sort through the massive volumes of state data generated by location and sensor technology. Zhang et al. (2014) presented an RFID simulation system to evaluate RFID deployment effectiveness in difficult environments, utilizing regression trees. Yi (2014) gave a graphic-based data mining approach to select useful data from a massive RFID data stream. Zhu (2014) also gave a graphical mining approach to track objects over a supply chain.

## 4.2. Data Mining Methodologies

There are widespread uses of data mining to support business (Olson and Shi 2006). Supply chain operations have been supported by the usual business data mining analyses of classification (to include customer profiling and fraud detection) and prediction. That implies use of the standard data mining methodologies of logistic regression, decision trees, and neural networks for classification, and regression for prediction. Clustering was applied by the National Iranian Oil Company Training Center (Hadighi et al. 2013) to integrate management by cluster in place of management by objectives. Costa et al. (2014) also reported us of cluster analysis in supply chains to monitor vehicle traffic management with data obtained from mobile devices. Genetic algorithms are a specialty tool that can be utilized in some data mining analyses. Ayoub et al. (2007) applied fuzzy clustering analysis and decision trees, using genetic algorithms to optimize location of biomass collection plants in bioenergy production in Japan.

Association rules provide a means to initially sort out relationships. Fuzzy analysis recognizes that some data points (especially subjective input) involves some range of uncertainty. Jain et al. (2008) used fuzzy association rule mining to help evaluate multiple criteria in agile supply chain decisions. Lau et al. (2009) also applied fuzzy association rules to capture logics operation data to identify root-causes of quality problems. Shih (2007) proposed use of multi-criteria field service scheduling in e-service supply chain management, while Ren et al. (2014) applied fuzzy classification to evaluate supply chain competitiveness.

Another technology often applied is the use of automation in the form of agents. This is especially attractive when dealing with big data, attempting to automate data gathering and analysis. Warkentin et al. (2012) suggested opportunities to utilize knowledge discovery through data mining in Web-based electronic commerce. Balasubramaniam and Thigarasu (2015) used agents in generating decision tree analysis to avoid over-fitting when measuring supply chain efficiency with respect to both cost and customer satisfaction.

## 4.3. Text Mining Papers

Data mining extends beyond mere numerical analysis. Text mining extends knowledge management to language data. Moriizumi et al. (2011) gave an example application for supply chain risk analysis, drawing upon academic research papers obtained from EBSCO and other search engines. Other papers report use of data mining tools such as support vector machines and random forests in text mining green supply chain management literature (Wu and Guo 2012). Sentiment analysis has grown in popularity. Ludwig et al. (2013) gave an example of analyzing Amazon.com book reviews. Schniederjans et al. (2013) analyzed negative word-of-mouth

information obtained from social media and their impact on financial performance. Ordenes et al. (2014) addressed noted limitations of text mining depth in customer feedback, suggesting an improved methodology.

Table 2. Recapitulation

| Section | Functions | 97 total |
|---|---|---|
| **Knowledge management** | Conceptual | 60 |
| | Impact on SCM strategy | |
| | Impact on SCM effectiveness | |
| | KM process support to SCM | |
| **Big Data** | Sources | 7 |
| | Linkages to SCM participants enabling collaboration | |
| | Impact on SMC planning | |
| | Enabler of mass customization | |
| | Enabler of risk mitigation | |
| **SCM functions** | | |
| Forecasting | Reduction of bullwhip effect | 18 |
| | Identification of cyclic activity | |
| Product development | Analysis of mass customization appropriateness | 30 |
| | Customer profiling | |
| | Mapping products to demands | |
| | Customer relationship management | |
| Supplier selection | Supplier selective via machine learning | 19 |
| | Multiple criteria analysis | |
| | Fuzzy analysis to reflect uncertain data | |
| Inventory management | Demand forecasting | 38 |
| | Association rules to model demand | |
| | Systems modeling to manage inventories | |
| Logistics | Intermodal transportation support | 47 |
| | Transportation network analysis | |
| | Transportation risk analysis | |
| Quality management | Quality function deployment | 33 |
| | Customer feedback analysis | |
| Risk analysis | Data security management | 21 |
| | Fraud detection | |
| **Tools** | | |
| RFID | RFID technology potential (and cost) | 22 |
| | Data exchange facilitation | |
| | Information technology support required | |
| | Data appropriateness analysis | |
| | Materiel tracking | |
| Data mining methodology | Data mining methodologies for classification | 74 |
| | Cluster analysis | |
| | Genetic algorithms | |
| | Decision tree algorithms | |
| | Association rules | |
| | Agent-based modeling | |
| Text mining | Sentiment analysis | 12 |
| | Word-of-mouth impact | |
| | Customer feedback | |

## 5. Discussion and Conclusion

Knowledge management in general has received a great deal of attention both in practice and in academic research. Technology continues to explode, and the ability to tap new sources of data has led to many interesting developments in the application of supply chain analytics. Technological developments such as cloud computing and RFID technology have enabled supply chain management to operate on a global basis across firms located around the globe.

Table 2 gives a recapitulation of what we have found. There are many papers in the data mining and forecasting literature for specific methodologies and applications. The papers reviewed here were the only ones selected by the search engine including both supply chain management and data mining, but a wealth of other related papers could be found by interested researchers. There were 97 papers on SCOPUS reflecting both "supply chain management" and "data mining." However, there were many papers with overlapping topics, as indicated by the last column of Table 2.

Table 2 demonstrates a wide variety of application areas found in this literature review. Data mining is a technical component of the field of knowledge management, and big data is a related topic focusing on the flood of data currently generated by our culture. This review establishes that all functions of supply chain management are supported by some aspect of data mining. The counts are not viewed as reflecting anything useful – simply which papers mentioned with term.  One cannot anticipate all future applications, so looking for gaps in coverage isn't really productive. Future innovations will come from understanding the potential of data mining and related technologies, utilizing cross-fertilization of ideas without constraint for any one view of classification. Future research will follow new innovations.

Data mining is not a cure-all for all supply chain operations. It has its limits. Data mining philosophy emphasizes the gaining of slight improvements through more responsive and accurate classification or prediction. There will always be errors, Type I and Type II. But the idea is to harness the wealth of data currently available in order to gain some slight advantage, often temporary and specific.

The future of data mining continues to look promising. Supply chains are here to stay, as they have been for millennia. Computer technology is much more recent. The application of this computer technology to generate measurements of important aspects of supply chains, and to analyze this data to make better decisions, will continue to grow.

# References

Ayoub, N., Martins, R., Wang, K., Seki, H. and Naka, Y. (2007) 'Two levels decision system for efficient planning implementation of bioenergy production', *Energy Conversion and Management,* 48(3), 709-723.

Balasubramaniam, C.P. and Thigarasu, V. (2015) 'Agent-based modeling in supply chain management using improved C4-5', *Research Journal of Applied Sciences, Engineering and Technology,* 9(2), 91-97.

Banerjee, A., Bandyopadhyay, T. and Acharya, P. (2013) 'Data analytics: Hyped up aspirations or true potential?', *Vikalpa: The Journal for Decision Makers,* 38(4), 1-11.

Bhattacharya, A., Kumar, S.A., Tiwari, M.K. and Talluri, S. (2014) 'An intermodal freight transport system for optimal supply chain logistics', *Transportation Research: Part C,* 38, 73-84.

Brandau, A. and Tolujevs, J. (2013)'Modelling and analysis of logistical state data', *Transport and Telecommunication,* 14(2), 102-115.

Chae, B. (2015) 'Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research', *International Journal of Production Economics,* 165(1), 247-259.

Chae, B., Olson, D. and Sheu, C. (2014) 'The impact of supply chain analytics on operational performance: A resource-based view', *International Journal of Production Research,* 52(16), 4695-4710.

Chen, M.C. and Wu, H.P. (2005) 'An association-based clustering approach to order batching considering customer demand patterns', *Omega,* 33(4), 333-343.

Chen, Y.S., Chen, C.H. and Lai, C.J. (2012) 'Extracting performance rules of suppliers in the manufacturing industry: An empirical study', *Journal of Intelligent Manufacturing,* 23(5), 2037-2045.

Choudhary, A.K., Harding, J.A., Lin, H.K., Tiwari, M.K. and Shankar, R. (2011) 'Knowledge discovery and data mining integrated (KOATING) Moderators for collaborative projects', *International Journal of Production Research,* 49(23), 7029-7057.

Collins, J.D., Worthington, W.J., Reyes, P.M.. and Romero, M. (2010) 'Knowledge management, supply chain technologies, and firm performance', *Management Research Review,* 33(10), 947-960.

Costa, G., Manco, G. and Masciari, E. (2014) 'Dealing with trajectory streams by clustering and mathematical transforms', *Journal of Intelligent Information Systems,* 42(1), 155-177.

Davenport, T. and O'Dwyer, J. (2011) 'Tap into the power of analytics', *Supply Chain Quarterly,* Fourth Quarter, 28-31.

Davis-Sramek, B., Germain, R. and Iyer, K. (2010) 'Supply chain technology: The role of environment in predicting performance', *Journal of the Academy of Marketing Science,* 38(1), 42-55.

Delen, D., Erraguntla, M., Mayer, R.J. and Wu, C.N. (2011) 'Better management of blood supply-chain with GIS-based analytics', *Annals of Operations Research,* 185(1), 181-193.

Demirkan, H. and Delen, D. (2012) 'Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud', *Decision Support Systems,* 55(1), 412-421.

Dobbs, R., Manyika, J. and Woetzel, J. (2015) *No Ordinary Disruption: The Four Global Forces Breaking All the Trends.* Philadelphia, PA, USA: PublicAffairs.

Faezy Razi, F. (2014) 'A supplier selection using a hybrid grey based hierarchical clustering and artificial bee colony', *Decision Science Letters,* 3(3), 259-268.

Forslund, H. and Jonsson, P. (2007) 'The impact of forecast information quality on supply chain performance', *International Journal of Operations & Production Management,* 27(1), 90-107.

Ghadge, A., Dani, S. and Kalawsky, R. (2012) 'Supply chain risk management: Present and future scope', *International Journal of Logistics Management,* 23(3), 313-339.

Gonzalez, H., Han, J., Cheng, H., Li, X., Klabjan, D. and Wu, T. (2010) 'Modeling massive RFID data sets: A gateway-based movement graph approach', *IEEE Transactions on Knowledge & Data Engineering,* 22(1), 90-104.

Ha, S.H. and Krishnan, R. (2008) 'A hybrid approach to supplier selection for the maintenance of a competitive supply chain', *Expert Systems with Applications,* 34(2), 1303-1311.

Hadighi, S.A., Sahebjamnia, N., Mahdavi, I. and Akbarpour Shirazi, M. (2013) 'A framework for strategy formulation based on clustering approach: A case study in a corporate organization', *Knowledge-Based Systems,* 49, 37-49.

Handfield, R. (2006) *Supply Market Intelligence: A Managerial Handbook for Building Sourcing Strategies.* New York: Taylor & Francis.

Handfield, R. and Nichols, E. (2004) 'Key issues in global supply base management', *Industrial Marketing Management,* 33(1), 29-35.

Ho, G.T.S., Lau, H.C.W., Kwok, S.K., Lee, C.K.M. and Ho, W. (2009) 'Development of a co-operative distributed process mining system for quality assurance', *International Journal of Production Research,* 47(4), 883-918.

Hong, G.H. and Ha, S.H. (2008) 'Evaluating supply partner's capability for seasonal products using machine learning techniques', *Computers and Industrial Engineering,* 54(4), 721-736.

Ivanov, D. and Sokolov, B. (2012) 'The inter-disciplinary modelling of supply chains in the context of collaborative multi-structural cyber-physical networks', *Journal of Manufacturing Technology Management,* 23(8), 976-997.

Jain, R., Singh, A.R., Yadav, H.C. and Mishra, P.K. (2014) 'Using data mining synergies for evaluating criteria at pre-qualification stage of supplier selection', *Journal of Intelligent Manufacturing,* 25(1), 165-175.

Jain, V., Benyoucef, L. and Deshmukh, S.G. (2008) 'A new approach for evaluating agility in supply chains using fuzzy association rules mining', *Engineering Applications of Artificial Intelligence,* 21(3), 367-385.

Jain, V., Wadhwa, S. and Deshmukh, S.G. (2009) 'Select supplier-related issues in modelling a dynamic supply chain: Potential, challenges and direction for future research', *International Journal of Production Research,* 47(11), 3013-3039.

Ji, S.W., Tian, Y. and Gao, Y.H. (2013) 'Study on supply chain information control tower system', *Information Technology Journal,* 12(24), 8488-8493.

Kohli, R. and Grover, V. (2008) 'Business value of it: An essay on expanding research directions to keep up with the times', *Journal of Association for Information Systems,* 9(1), 23-39.

Kraus, C. and Valverde, R. (2014) 'A data warehouse design for the detection of fraud in the supply chain by using the Benford's law', *American Journal of Applied Sciences,* 11(9), 1507-1518.

Lau, H.C.W., Ho, G.T.S., Zhao, Y. and Chung, N.S.H. (2009) 'Development of a process mining system for supporting knowledge discovery in a supply chain network', *International Journal of Production Economics,* 122(1), 176-187.

Lavalle, S., Lesser, E., Shockey, R.H., M and Kruschwitz, N. (2011) 'Big data, analytics and the path from insights to value', *MIT Sloan Management Review,* 52(2), 21-32.

Le, H.Q., Arch-Int., S., Nguyen, H.C. and Arch-Int., N. (2013) 'Association rule hiding in risk management for retail supply chain collaboration', *Computers in Industry,* 64(7), 776-784.

Lee, C.K.M., Lau, H.C.W., Kwok, S.K. and Ho, G.T.S. (2010) 'Design and development of supply chain knowledge discovery system for customer relationship management', *International Journal of Services, Technology and Management,* 14(1), 2-16.

Li, G.D., Yamaguchi, D. and Nagai, M. (2008) 'A grey-based rough decision-making approach to supplier selection', *International Journal of Advanced Manufacturing Technology,* 36(9-10), 1032-1040.

Li, Y., Kramer, M.R., Beulens, A.J.M. and van der Vorst, J.G.A.J. (2010) 'A framework for early warning and proactive control systems in food supply chain networks', *Computers in Industry,* 61(9), 852-862.

Liao, S.H., Chen, Y.N. and Tseng, Y.Y. (2009) 'Mining demand chain knowledge of life insurance market for new product development', *Expert Systems with Applications,* 36(5), 9422-9437.

Liao, S.H., Chen, C.M. and Wu, C.H. (2008a) 'Mining customer knowledge for product line and brand extension in retailing', *Expert Systems with Applications,* 34(3), 1763-1776.

Liao, S.H., Hsieh, C.L. and Huang, S.P. (2008b) 'Mining product maps for new product development', *Expert Systems with Applications,* 34(1), 50-62.

Liao, S.H. and Wen, C.H. (2009) 'Mining demand chain knowledge for new product development and marketing', *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews,* 39(2), 223-227

Liberatore, M. and Luo, W. (2010) 'The analytics movement', *Interfaces,* 40(4), 313-324.

Lin, R.H., Chuang, C.L., Liou, J.J.H. and Wu, G.D. (2009) 'An integrated method for finding key suppliers in SCM', *Expert Systems with Applications,* 36(3) part 2, 6461-6465.

Ludwig, S., de Ruyter, K., Friedman, M., Brűggen, E.C., Wetzels, M. and Pfann, G. (2013) 'More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates', *Journal of Marketing,* 77(1), 87-103.

Masciari, E. (2012) 'SMART: Stream monitoring enterprise activities by RFID Tags', *Information Sciences* 195, 25-44.

Min, H. (2006) 'Developing the profiles of supermarket customers through data mining', *Service Industries,* 26(7), 747-763.

Mithas, S., Ramasubbu, N. and Sambamurthy, V. (2011) 'How information management capability influences firm performance', *MIS Quarterly,* 35(1), 237-256.

Morizumi, S., Chu, B., Cao, H. and Matsukawa, H. (2011) 'Supply chain risk driver extraction using text mining technique', *Information,* 14(6), 1935-1945.

Ni, M., Xu, X. and Deng, S. (2007) 'Extended QFD and data-mining-based methods for supplier selection in mass customization'. *International Journal of Computer Integrated Manufacturing,* 20(2-3), 280-291.

O'Dwyer, J. and Renner, R. (2011) 'The promise of advanced supply chain analytics', *Supply Chain Management Review,* 15, 32-37.

Oliva, R. and Watson, N. (2011) 'Cross functional alignment in supply chain planning: A case study of sales & operations planning', *Journal of Operations Management,* 29(5), 434-448.

Olson, D.L. and Shi, Y. (2006) *Introduction to Business Data Mining*. Irwin: McGraw-Hill.

Ordenes, F.V., Theodoulidis, B., Burton, J., Gruber, T. and Zaki, M. (2014) 'Analyzing customer experience feedback using text mining: A linguistics-based approach', *Journal of Service Research,* 17(3), 278-295.

Qin, L. and Zhao, X. (2012) 'Design and realization of information service platform of logistics parks based on cloud computing', *Advances in Information Sciences and Service Sciences,* 4(23), 112-120.

Parmar, D., Wu, T., Callarman, T., Fowler, J. and Wolfe, P. (2010) 'A clustering algorithm for supplier base management', *International Journal of Production Research,* 48(13), 3803-3821.

Peng, M., Peng, Y. and Chen, H. (2014) 'Post-seismic supply chain risk management: A system dynamics disruption analysis approach for inventory and logistics planning', *Computers and Operations Research,* 42, 14-24.

Pillutla, S., Yao, D.Q. and Li, X. (2014) 'Data mining the adoption intention of e-procurement system for Chinese companies', *International Journal of Management & Decision Making,* 13(2), 207-219.

Piramuthu, S. (2005) 'Knowledge-based framework for automated dynamic supply chain configuration', *European Journal of Operational Research,* 165(1), 219-230.

Ren, C., Liu, Y. and Guo, Y. (2014) 'Fuzzy evaluation on supply chain competitiveness based on membership degree transformation new algorithm', *Journal of Chemical and Pharmaceutical Research,* 6(2), 139-144.

Renaud-Gentié, C., Burgos, S. and Benoît, M. (2014) 'Choosing the most representative technical management routes within diverse management practices: Application to vineyards in the Loire Valley for environmental and quality assessment', *European Journal of Agronomy,* 56, 19-36.

Robinson, A., Levis, J. and Bennet, G. (2010) 'INFORMS to officially join analytics movement', *OR/MS Today,* 37(5), 59.

Sathi, A. (2012) *Big Data Analytics: Disruptive Technologies for Changing the Game*. Boise, ID: MC Press.

Schlegel, G.L. (2014/2015) 'Utilizing big data and predictive analytics to manage supply chain risk', *Journal of Business Forecasting,* 33(4), 11-17.

Schniederjans, D., Cao, E.S. and Schiederjans, M. (2013) 'Enhancing financial performance with social media: An impression management perspective', *Decision Support Systems,* 55(4), 911-918.

Schoenherr, T. and Speier-Pero, C. (2015) 'Data science, predictive analytics, and big data in supply chain management: Current state and future potential', *Journal of Business Logistics,* 36(1), 120-132.

Shao, C., Wang, L., Xiao, L. and Wu, J. (2010) 'Qualitative phase space reconstruction analysis of supply-chain inventory time series', *South African Journal of Science,* 106(11-12), 1-7.

Shih, S.C. (2007) 'A three-stage field service management model for effective post-sales service supply chain management', *International Journal of Manufacturing Technology and Management,* 12(4), 384-404.

Shukla, V., Naim, M.M. and Thornhill, N.F. (2012) 'Rogue seasonality detection in supply chains', *International Journal of Production Economics,* 138(2), 254-272.

Smith, A.D. (2011) 'Managing supply chain complexities and integration issues for competitive advantage: A comparative analysis', *International Journal of Logistics Systems and Management,* 9(3), 351-374.

Soeanu, A., Debbabi, M., Alhadidi, D., Makkawi, M., Allouche, M., Bélanger, M. and Léchevin, M. (2015) 'Transportation risk analysis using probabilistic model checking', *Expert Systems with Applications,* 42(9), 4410-4421.

Song, D.P. and Dinwoodie, J. (2008) 'Quantifying the effectiveness of VMI and integrated inventory management in a supply chain with uncertain lead-times and uncertain demands', *Production Planning and Control,* 19(6), 590-600.

Song, Z. and Kusiak, A. (2009) 'Optimising product configurations with a data-mining approach', *International Journal of Production Research,* 47(7), 1733-1751.

Srivihok, A. and Intrapairot, A. (2014) 'Model of inbound ASEAN economics community tourists in Thailand by using twining of feature selection and classification algorithms', *Advanced Science Letters,* 20(10-12), 2202-2205.

Stefanovic, N. (2014) 'Proactive supply chain performance management with predictive analytics', *Scientific World Journal,* 2014*,* 1-17.

Stefanovic, N. (2015) 'Collaborative predictive business intelligence model for spare parts inventory replenishment', *Computer Science and Information Systems,* 12(3), 911-930.

Thiesse, F., Floerkemeier, C., Harrison, M., Michahelles, F. and Roduner, C. (2009) 'Technology, standards, and real-world deployments of the EPC network', *IEEE Internet Computing,* 13(2), 36-43.

Thomassey, S. (2010) Sales forecasts in clothing industry: 'The key success factor of the supply chain management', *International Journal of Production Economics,* 128(2), 470-483.

Tien, J.M. (2006) 'Data mining requirements for customized goods and services', *International Journal of Information technology & Decision Making,* 5(4), 683-698.

Tien, J.M. (2012) 'The next industrial revolution: Integrated services and goods', *Journal of Systems Science and Systems Engineering,* 21(3), 257-296.

Ting, S.L., Tse, Y.K., Ho, G.T.S., Chung, S.H. and Pang, G. (2014) 'Mining logistics data to assure the quality in a sustainable food supply chain: A case in the red wine industry', *International Journal of Production Economics,* 152, 200-209.

Tirkel, I. (2013) 'Forecasting flow time in semiconductor manufacturing using knowledge discovery in databases', *International Journal of Production Research,* 51(18), 5536-5548.

Tsou, C.M. (2013) 'On the strategy of supply chain collaboration based on dynamic inventory target level management: A theory of constraint perspective', *Applied Mathematical Modelling,* 37(7), 5204-5214.

Wang, C., Luxhøj, J.T. and Johansen, J. (2004) 'Applying a knowledge management modeling tool for manufacturing vision (MV) development', *Industrial Management & Data Systems,* 104(9), 735-743.

Warkentin, M., Sugumaran, V. and Sainsbury, R. (2012) 'The role of intelligent agents and data mining in electronic partnership management', *Expert Systems with Applications,* 39(18), 13277-13288.

Wu, S. and Guo, J. (2012) 'The text mining and classification analyses on the relationship between green supply chain management and closed-loop supply chain research trends (2000-2010)', *International Journal of Digital Content Technology and its Applications,* 6(23), 281-288.

Xu, X. and Lin, J. (2009) 'An integrated method to discovering key suppliers in supplier chain management', *Journal of Computational Information Systems,* 5(3), 1445-1452.

Yi, L. (2014) 'A graphic-based data mining approach in RFID sensor networks', *International Journal of Control and Automation,* 7(2), 387-396.

Yűksel, M.E. and Yűksel, A.S. (2011) 'RFID technology in business systems and supply chain management', *Journal of Economic & Social Studies,* 1(1), 53-71.

Zhang, T., Yin, Y., Yue, D., Wang, X. and Yu, G. (2014) 'Research and implementation of an RFID simulation system supporting trajectory analysis', *Journal of Software,* 9(1), 162-168.

Zhao, K. and Yu, X. (2011) 'A case based reasoning approach on supplier selection in petroleum enterprises', *Expert Systems with Applications,* 38(6), 6839-6847.

Zhou, M., Dresner, M. and Windle, R.J. (2008) 'Online reputation systems: Design and strategic practices', *Decision Support Systems,* 44(4), 785-797.

Zhu, X. (2014) 'An algorithm for moving path discovery using frequent graphical mining approach', *International Journal of Control and Automation,* 7(10), 145-154.

Zou, Z., Tseng, T.L., Sohn, H., Song, G. and Gutierrez, R. (2011) 'A rough set based approach to distributor selection in supply chain management', *Expert Systems with Applications,* 38(1), 106-115.