2016-04

The impact of big data on world-class sustainable manufacturing

Dubey, R

http://hdl.handle.net/10026.1/5175

10.1007/s00170-015-7674-1
The International Journal of Advanced Manufacturing Technology
Springer Science and Business Media LLC

The impact of Big Data on World Class Sustainable Manufacturing

Abstract

Big data (BD) has attracted increasing attention from both academics and practitioners. This paper aims at illustrating the role of Big Data analytics in supporting world-class sustainable manufacturing (WCSM). Using an extensive literature review to identify different factors that enable the achievement of WCSM through BD and 405 usable responses from senior managers gathered through social networking sites (SNS), we propose a conceptual framework that summarizes this role, test this framework using data which is heterogeneous, diverse, voluminous, and possess high velocity, and highlight the importance for academia and practice. Finally we conclude our research findings and further outlined future research directions.

Key words: Big Data, World Class Sustainable Manufacturing, Social Networking Site, Confirmatory factor Analysis, Sustainable Manufacturing.

1. Introduction

In recent years Big Data Analytics (BDA) has been an important subject of debate among academics and practitioners. McKinsey Global Institute has predicted that by 2018 the BDA needs for the United States alone will be more than 1.5 million managers who need to possess skills in analyzing Big Data for effective decision making. In developing countries, in the recent 13th Confederation of Indian Industries manufacturing summit, BDA was at the forefront of discussions among manufacturing professionals in India. The Internet of things (IOT) and big data & predictive analytics are now within the reach of the operations management community to begin to explore, with the potential for measurable and meaningful impacts on the life of people in the developing world (Accenture, 2013). On the other hand, thinkers such as Professor Nassim Nicholas Taleb, in his interview in the Economic Times highlighted the impacts of BD, but was skeptical about its success.

The literature on the role of BDA in Operations and Supply Chain Management (OM/SCM) (for example Wamba et al., 2015) has argued for benefits from its use, including, *inter alia*, 15-20% increase in ROI (Perrey et al., 2013), productivity and competitiveness for companies and public sector, as well as economic surplus for customers (Manyika et al., 2011), and informed decision making that allows visibility in operations and improved performance measurement (McAfee and Brynjolfsson, 2012).

The majority of studies so far have endeavored to understand the different dimensions of the concept and to capture the potential benefits to OM/SCM (Chen et al., 2013; Wamba et al., 2015). There is little known about the contribution of BDA to sustainability practices, and in particular the role of BDA in achieving world class sustainable manufacturing, especially from a developing countries perspective. "World-class manufacturing" (WCM) was coined by Hayes and Wheelwright (1984) to denote "*a set of practices, implying that the use of best practices would lead to superior performance. This practice-based approach to world class manufacturing has been echoed by numerous authors since then*"… (Flynn et al. 1999). In our study, world-class sustainable manufacturing (WCSM) is defined as that set of practices that would lead to superior sustainability performance. Keeso (2014), in his recent review of the role of BDA for sustainability, suggests that "big data adoption has broadly been slow to coalesce with sustainability efforts" (p.2), but still he has focused on BDA and the environmental aspect of sustainability. In the present paper our contribution is largely restricted to "big data and analytics" (BDA) in extending the literature on WCSM and understanding how in future big data can be exploited in other fields.

Driven by the need to further explore the role of BDA for WCSM, this paper acts to bridge this knowledge gap by achieving the following objectives: (i) to clarify the definition of BDA and its relationship to WCSM; (ii) to propose a conceptual framework that summarizes this role; (iii) to test the proposed sustainability framework using data which is heterogeneous, diverse, voluminous, and possesses high velocity; (iv) to develop future directions on the role of BDA in WCSM.

The paper is organized as follows. The next section reviews the literature on BDA and WCSM and identifies research gaps. In the third section, we will focus on model development, whereas the fourth section focuses on research design. The fifth and sixth sections present the psychometric properties of the measuring items (i.e. reliability and validity of constructs) and findings. Finally, the paper discusses the contribution to the literature, the limitations of the work, and outlines further research directions.

2. Literature Review

2.1 Big Data

'Big Data and Analytics' (BDA) has attracted the attention of scholars from every field including, genomics, neuroscience, economics and finance (Fan et al. 2014). BDA is one of the fastest evolving fields due to convergence of internet of things (IOT), the cloud and smart assets (Bughin et al. 2010). Mayer-Schonberger and Cukier (2013) have argued that there is no rigorous definition of "big data". Manyika et al. (2011) have argued that BD is the next frontier for innovation that may provide competitive advantage to organizations. In this paper, we follow Dijcks (2013) with the definition of BD as: (i) traditional enterprise data, machine generated, or data stemming from weblogs, sensors and logs, and (ii) social data. Since there is a mass of information generated from this data, this raises challenges for organizations with regard to data storage, analysis and processing, and value, as well as concerns regarding the security and ownership. BD is characterized by (i) volume, denoting the large amount of data that need to be stored or the large number of records; (ii) velocity, denoting the frequency or speed by which data is generated and delivered; and (iii) variety, which illustrates the different sources by which data is generated, either in a structured or unstructured format (Wamba et al., 2015). White (2012) has added the fourth dimension, veracity, to highlight the importance of quality data and the level of trust in a data source. Besides the four characteristics, scholars (e.g. Forrester, 2012) have also added another dimension, value, to denote the economic benefits from the data.

In this research, we echo the views of Wamba and colleagues as well as McAfee et al. (2012) and focus on the four main dimensions of BD. This is because these characteristics affect decision-making behaviours, and also create critical challenges. Boyd and Crawford (2012) have argued that big data is a cultural, technological, and scholarly phenomenon that revolves around technology, analysis, and mythology. According to Mark and Douglas (2012), BD is defined as high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information for enhanced insight and decision making. McGahan (2013) further argues that big data is too large to handle with conventional software programs such as Excel, and thus requires specialized analytics. Sun et al. (2015) have argued that big data is data whose sources are heterogeneous and autonomous; whose dimensions are diverse; whose size is beyond the capacity of conventional processes or tools to effectively and affordably capture, store, manage, analyze, and exploit; and whose relationships are complex, dynamic, and evolving.

Gandomi and Haider (2015) have attempted to further our understanding of BD and of its potential applications. While the majority of the literature is focussed more on BD technology and predictive analytics, Gandomi amd Haider (2015) have attempted to provide detailed explanations for volume, variety, velocity, veracity, variability and value. In the same work they have outlined various techniques and tools that can enhance decision making abilities that were limited during the traditional data era (i.e. text analytics, audio analytics, video analytics, social media analytics, and predictive analytics). Some scholars may focus on the variety dimension (Davenport et al., 2012) while others emphasise the importance of storage and analysis (Jacobs, 2009; Manyika et al., 2011) highlighting the role of analytics. This role is further explicated in the next section.

2.2 Big Data Analytics and applications in Operations and Supply Chain Management

Waller and Fawcett (2013) underline the importance of data and analytics for SCM. They introduce the term 'SCM data science', referring to BDA, as the "application of quantitative and qualitative methods from a variety of disciplines in combination with SCM theory to solve relevant SCM problems and predict outcomes, taking into account data quality and availability issues" (p. 79). Bi and Cochran (2014) argue that BDA has been identified as a critical technology to support data acquisition, storage, and analytics in data management systems in modern manufacturing. They attempt to connect IOT and BD to advanced manufacturing information systems to help to streamline the existing bottlenecks through improving forecasting systems. Similarly, Gong et al. (2014) argue that a production control system (PCS) can be considered an information-processing organization (IPO). They conclude that the existing literature surrounding PCS has not given attention to decision-making efficiency. Thus the delay in information generation through analysis may hamper the performance of the production systems. The use of BDA can further streamline the data bottlenecks that currently plague the performance of MRP, KANBAN, and CONWIP. Hazens et al. (2014) have argued that supply chain professionals are inundated with data, motivating new ways of thinking about how data are produced, organized, and analyzed. Hence the volume, variety and velocity of data provide impetus to the organizations to adopt and perfect data analytic functions (e.g. data science, predictive analytics, and big data) to improve the current supply chain processes and their performance. In the article the authors have clearly argued the need for quality data to examine the current supply chain processes using organizational theories. Chae (2015) has argued that in the present situation, social media and big data are complementary to each other. Chae (2015) have further noted that the field of operations management has been relatively slow in studying BD and social media. The author proposes a conceptual framework related to use of Twitter to understand current trends in SCM. Li et al. (2015) have discussed the potential application of big data in product life cycle management. However, the implications of BDA for world-class manufacturing (WCM) and its extension from a sustainability point of view (i.e. World class sustainable manufacturing) have not yet been realized. We discuss WCM and WCSM in the next section.

2.3 World-Class Manufacturing

World-class manufacturing (WCM) was first introduced by Hayes and Wheelwright (1984) (see Flynn et al. 1999). Hayes and Wheelwright (1984) have related WCM to those practices that aim at enabling superior performance (Flynn et al. 1999). Since 1986, Schonberger's work on WCM has attracted major attention from academia and practitioners. He argued that those manufacturing organizations that have consistently performed in terms of superior market performance have embraced five common practices - just-in-time (JIT), total quality management (TQM), total productive maintenance (TPM), employee involvement (EI) and simplicity. Hall (1987) has further identified common practices among world class manufacturing organizations as total quality, JIT and people involvement. Gunn (1987) identified world class manufacturing practices as total quality, supplier relations, customer focus, lean manufacturing/operations, computer integrated manufacturing and distribution and services after sales. Steudel and Desruelle (1992) identified practices that separate world class manufacturers from traditional manufacturing organizations - total quality, supplier relationship, employee involvement, lean operations, total productive maintenance and group technology. According to Roth et al. (1992) employee involvement, manufacturing strategy and vision, innovation, and performance measurement are the practices that make a manufacturing organization a "world class manufacturing" organization. Flynn et al. (1997) have outlined that top management commitment, customer relationship, supplier relationship, work force management, work attitudes, product design process, statistical control and feedback, and process-flow management are the some of the practices which explain the consistent performance of the manufacturing organizations. Brown et al. (2007) have identified that employee involvement, manufacturing strategy and business strategy separate world class manufacturing organizations from traditional manufacturing organizations. Sharma and Kodali (2008) have identified practices of WCM as manufacturing strategy, leadership, environmental manufacturing, human resource management, flexible management, supply chain management, customer relationship management, production planning, total quality management, total productive maintenance and lean manufacturing.

The focus of WCM on customer satisfaction through satisfying the appropriate performance objectives (speed, flexibility, dependability, quality, cost) suggest the importance of acquiring, storing, and analyzing BD for, inter alia, decision making, innovation, visibility, customization of products and services, and ultimately sustainable competitive advantage (Wamba et al., 2015). Furthermore, mirroring the need expressed by organizations to achieve superior performance but considering at the same time the environmental and social consequences of their endeavors, we highlight the importance of BD for sustainable WCM, which is discussed in the next section.

2.4 Sustainable Manufacturing Practices

Sustainable manufacturing is a strategy of development of new products. It is defined by the U.S. Department of Commerce (2007) as “the creation of manufactured products that use processes that minimize negative environmental impacts, conserve energy and natural resources, are safe for employees, communities, and consumers and are economically sound.” The integration of environmental requirements throughout the entire lifetime of product needs a new way of thinking and new decision tools to be applied (Kaebernick et al. 2003; Jovane et al. 2008; Garetti and Taisch, 2012). Thus sustainable manufacturing involves green product design, green procurement, green technology and green production (Noci, 1997; Azzone and Noci, 1998; Gunasekaran and Spalanzani, 2012). Manufacturing practices have evolved over the last two decades from traditional manufacturing, concerned with cost, quality, delivery and flexibility (Sanchez and Perez, 2001) to sustainable manufacturing which aims at achieving a balance between environmental, social and economic dimensions to satisfy stakeholders (Flammer, 2013) and achieve competitive advantage (Rusinko, 2007; Carter and Rogers, 2008; Kannegiesser and Gunther, 2014). Molamohamadi and Ismail (2013) have argued that technology, education, ethnic background and accountability are the key enablers of sustainable manufacturing. Prabhu et al. (2012) have argued that the minimization of energy consumption and waste minimization are two key aspects of sustainable manufacturing. Gunasekaran et al. (2013) have argued that operational strategies, tactics & techniques and operational policies are the foundation of sustainable manufacturing. Garbie (2013, 2014) has further argued that to implement sustainable manufacturing, an organization needs to focus on key enablers such as international issues, contemporary issues, innovative products, reconfigurable manufacturing systems, complexity analysis, lean production, agile manufacturing, performance measurement and flexible organization. Dubey et al. (2015) have further attempted to take the sustainable manufacturing practices to world-class sustainable manufacturing level. The pillars identified are leadership, regulatory pressures, supplier relationship management, employee involvement, reconfigurable manufacturing systems, lean production, and agile manufacturing..

Literature has discussed sustainable manufacturing (e.g. Lovins et al., 1999) and sustainable practices such as waste minimization and energy efficiency through monitoring or technology (Despeisse et al., 2013). However, to be able to implement sustainable manufacturing and achieve superior performance by excelling in the three pillars of sustainability performance, that is, economic, environmental, and social, organizations need to make use of large amounts of data, that is, BD. Organizations need to acquire, store, analyze, and use BD in order to take decisions related to the achievement of their supply chain and strategy goals. Therefore, there is need for BDA adoption within WSSCM. Garetti and Taisch (2012), in their review of sustainable manufacturing, highlight the role of data and BDA, suggesting that there is need for methods that will be able to process large amounts of data related to environmental, social, and economic implications. BDA is therefore needed within WCSM.

2.5 Research Gap

Despite the growing interest in WCSM, there is still lack of consensus in current literature with regards to its definition and implication for organizations (Garetti and Taisch, 2012). Additionally, the majority of research has explored issues such as performance, operational strategies and techniques to achieve competitive advantage (Rusinko, 2007; Kannegiesser and Gunther, 2014; Dubey et al., 2015). Although the aforementioned scholars recognize the need for BDA within WCSM, there is yet research to be conducted to address the role of BDA. Current studies (e.g. Opresnik and Taisch, 2015) have investigated how manufacturers could harness the benefits of BDA for servitization, suggesting that BD are vital to this process. However, they have mainly focused on 'value' and not on volume, velocity, and variety. They also do not focus on the role of BD on WCSM. We aim to address these gaps and are driven by the endorsement of the European Commission on Industrial Technologies Research to study sustainable manufacturing not only in Europe, but also on a global level to address the challenges related (Garetti and Taisch, 2012).

3. Theoretical Framework

We propose a framework to investigate the importance of BDA for WCSM (see Figure 1). We have identified the constructs which impact upon sustainable manufacturing on the basis of extensive literature review followed by principal component analysis (PCA*) on the set of data collected (see Appendix 1). The foundations of our theoretical framework are grounded in the data we have gathered. In Figure 1 the constructs represented as X1, X2, X3, X4…., Xn represent orthogonal factors which we have derived using suitable data reduction methods as discussed in Section 5.2. We argue the constructs are formative and further they have reflective nature. Each of the constructs is studied from a BDA perspective, which is discussed in our research design section.

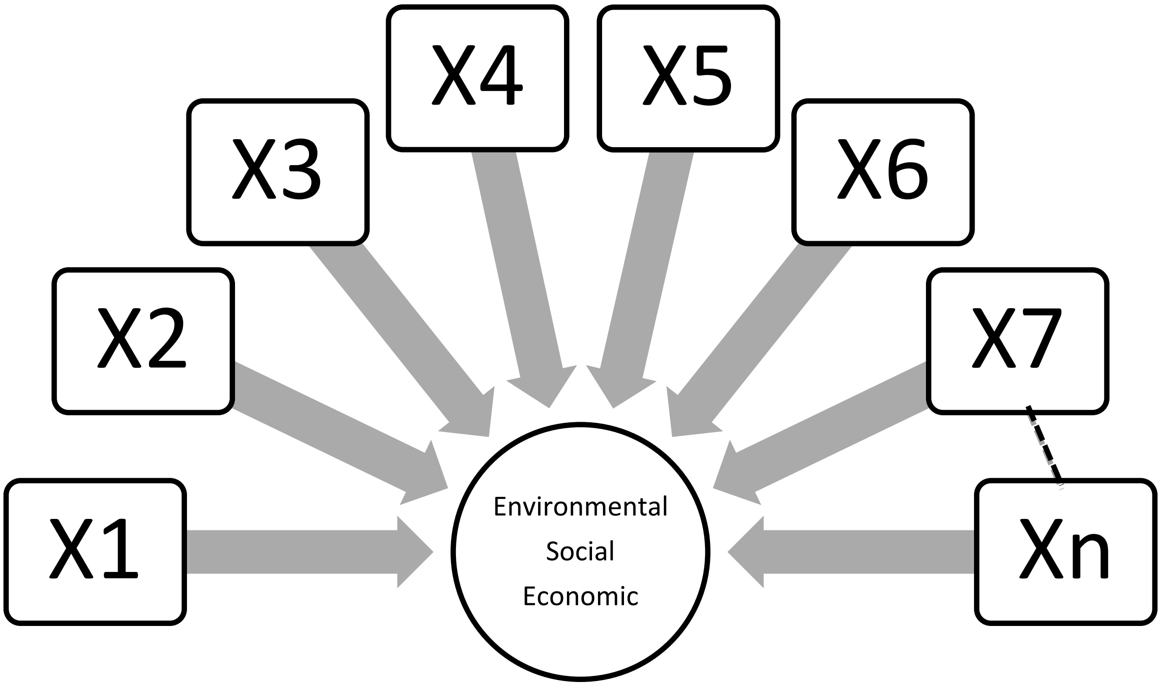

Figure 1: BDA and World Class Sustainable Manufacturing Framework

(Note: * *In our case we have transformed a (405 x 51) data matrix into (405 x 9) data matrix. Hence "n" is not that large, so the data matrix was easily reduced using PCA. However if "n" had been extremely large then we would have used "RP" for reduction as per discussion in our preceding section*)

3.1 Building Blocks of World Class Sustainable Manufacturing Framework

We explain each construct and their items of WCSM framework in tabulated form as shown in Table 1.

Table 1: Building blocks of WCSM framework and their indicators

Building Blocks	**Reference**	**Indicators**
Leadership	Siaminwe et al. (2005); Berkel (2007); Deif (2011); Despeisse et al. (2012); Law and Gunasekaran (2012); Singh et al. (2012); Dues et al. (2013); van Hoof and Lyon (2013); Dubey et al. (2015); Dutta and Bose (2015)	• Well defined environmental policy • Awareness about environmental policy • Top management support • Top management has approved special fund for investment in cleaner technologies • Top management positive attitude towards green practices Ssenior managers motivate and support new ideas received from junior executives • Recognition of employees
Regulatory Pressures	Zhu et al. (2005); Tsoulfas and Pappis (2006); Sarkis et al. (2011); Singh et al. (2012) ; Dubey et al. (2015)	• A regional pollution control board pressurizing the firm to adopt green practices; • Government regulations provide clear guidelines in controlling pollution level; • Pollution control board strictly monitors the pollution level of firms on a periodic basis; • Green practices decrease incidence of penalty fee charged by pollution control board
Supplier Relationship Management	Bierma and Waterstraat (1999); Vachon and Klassen (2006); Hsu and Hu (2009); Bai and Sarkis (2010); Ku et al. (2010); Testa and Iraldo (2010); van Hoof and Lyon (2013); Dubey et al. (2015)	• Environmental criteria considered while selecting suppliers; • Firm considers environment collaboration with suppliers; • Firm has technological integration with suppliers; • Firm trains and educates suppliers in implementing ISO 14001; • Environmental audit for suppliers

		done periodically
Employee involvement	Atlas and Florida (1998); Chien and Shih (2007); Hsu and Hu (2008); Luthra et al. (2011); Jabbour et al. (2013); Dutta and Bose (2015)	• Strategic participation; • Organizational participation; • Task discretion;
Customer Relationship	Rao and Holt (2005); Vachon and Klassen (2006); Seuring and Muller (2008); Eltayeb et al. (2011); Baines et al.(2012)	• Green practices improve customer satisfaction; • Firm recovers end of life products from customers; • Customers appreciate eco-friendly products;
Total Quality Management	Pauli (1997); Murovec et al. (2012); Prajogo et al. (2012); Pereira-Moliner et al. (2012); Gavronski et al. (2013)	• Involvement of top management; • Strategic quality management planning; • Customer focus / customer satisfaction; • Employee training for quality; • Supplier quality assurance and management; • Quality information management and analysis; • ISO 9000:2000; • TQM tools, techniques, systems and resources in place;
Total Productive Maintenance	Mudgal et al. (2010); Diaz-Elsayed et al. (2013); Jasiulewicz-Kaczmarek, (2013)	• Maintenance strategy and policy deployment ownership; • Process / equipment classification, standardization and improvement; • Process quality maintenance; • Maintenance practices/ procedures/ practices; • Standardization of materials, machines and methods (3M's);

Lean manufacturing	Farish (2009); Franchetti et al. (2009); Deif (2011); Dues et al. (2013); Hajmohammad et al. (2013); Garbie (2013, 2014)	• JIT tools, techniques and processes; • Standardized work/standard operations; • Cycle time/lead time/lot-size reduction • Cellular manufacturing/focused factory • Mixed model assembly/mass customization ; • Pull system;
Environment	Carter and Rogers (2008); Azevedo et al. (2011); Deif (2011); Bhateja, et al. (2012); Seman et al. (2012); Whitelock (2012)	• Environmental technology; • Recycling efficiency; • Eco packaging; • Level of process management which includes pollution control, waste emissions, carbon footprints etc;
Social	Carter and Rogers (2008); Pochampally et al. (2009); Gunasekaran and Spalanzani (2012); Dues et al. (2013); Gavronski et al. (2013)	• Management commitment; • Customer satisfaction • Employee development;
Economic	Carter and Rogers (2008); Azevedo et al. (2011); Ageron et al. (2012).	• Environmental cost; • Supply chain cost; • Cost to quality; • Responsiveness cost;

4. Research Design

4.1 Measures

Measures were adopted or modified from scales identified from extant literature to avoid scale proliferation. We used multi-item measures of constructs for our theoretical model in order to improve reliability, reduce measurement error, ensure greater variability among survey individuals, and improve validity (Churchill, 1979). Each construct was operationalized using at least three indicators for effective measurement and analysis, applying confirmatory factor analysis (Gerbing and Anderson, 1988). Table 3 summarizes the scales.

All indicators included in the survey were pretested to ensure precise operationalization of defined variables in the survey instrument.

4.2 Sampling Design

We identified large manufacturing firms that have more than 1000 employees and an annual turnover of more than 2 billion INR. The initial sample frame consisted of 1130 manufacturing firms and was compiled from databases provided by CII-Institute of Manufacturing.

4.3 Data Collection

Data was collected through social networking sites (SNS). Lomborg and Bechmann (2014) have argued that APIs (Application Programming Interfaces) can be very useful for collecting data from social media in an ethical manner. SNS have now become increasing important for data scientists (Hargittai, 2007). Prior to questioning, respondents were told that responses would be kept strictly confidential. We sent our questionnaire to those respondents who accepted our request on Facebook or LinkedIn to respond to our survey. In this way we could reach the maximum number of respondents within a few weeks in comparison to traditional methods such as e-mail, where respondents may not respond to the e-mail, or automatically delete it or render it spam. We included LinkedIn, Facebook and Twitter (see Berg et al. 2004; Tufekci, 2008; Kwak et al. 2010). They were chosen since response is comparatively fast (velocity) in comparison to traditional data collection procedures, variety was allowed (other details can be easily acquired which company reports do not provide), volume (large sample size can be reached within shortest time), veracity (through multiple accounts like Facebook, Twitter and LinkedIn) the authenticity of the information's can be easily checked which traditional data collection does not offer. Overall we received 280 complete and usable responses. We further followed up with other respondents and within a month we received another 125 complete and usable responses. In this way we received 405 complete and usable responses, which represent 35.84%. The response size is quite high in comparison to similar studies conducted in the OM/SCM field using traditional data collection methods (e.g. Braunscheidel and Suresh, 2009; Dubey et al. 2015). The demographic profile of the respondents is shown in Table 2.

Table 2: Demographic profile of the respondents

Designation		**Number of respondents**	**Percentage of respondents**
	Vice President	76	18.77
	General Managers	85	20.99
	Managers	110	27.16
	Deputy/Assistant Managers	134	33.09
Work experience (years)	Above 20	140	34.57
	15–20	35	8.64
	10–14	40	9.88
	5–9	85	20.99
	0-4	105	25.93
Type of business	Auto components manufacturing	135	33.33
	Heavy Machinery	45	11.11
	Electrical Components	37	9.14
	Infrastructure Sector	30	7.41
	Steel Sector	35	8.64
	Chemical	123	30.37
Age of the firm	>20	90	26.95
	15-20	220	46.11
	10–14	75	16.17
	5–9	20	10.77
	0-4	0	0

Revenue (Indian Rupees INR)	> 2000 crores	50	12.35
	1500-2000 crores	80	19.75
	1000-1499 crores	170	41.98
	500-999 crores	100	24.69
	< 500	5	1.23
Number of employees	Greater than 500	200	49.38
	250-500	150	37.04
	100-249	35	8.64
	Less than 100	20	4.94

From Table 2 we can see that around 40% of the respondents are in senior positions in their companies. This may explain why approaching the respondents through SNS may have better response rate in comparison to sending e-mail and following up several times for response. In recent years many companies have policies in place that do not encourage their employees to respond to questionnaires (Eckstein et al. 2015). The majority of responses gathered were from auto components manufacturing firms. These firms in India are quite responsible towards P's (planet, people, and profit).

5. Testing of Big Data

Fan et al. (2014) argued that big data possess unique properties. We have gathered data from SNS, hence our gathered data may possess high volume and variety but testing is required to address possible challenges during data analysis such as heterogeneity, noise accumulation, spurious correlation, and incidental endogeneity. We discuss their assessment in the next sections.

5.1.1 Heterogeneity

Big data results from data accumulation from various multiple sources corresponding to different subpopulations. Fan et al. (2014) have argued that these subpopulations may exhibit some different unique properties not shared by others. In case of traditional data sets where sample size is small or moderate, data points from small subpopulations are referred as outliers and these outliers may impact the final outcome of statistical analyses. However, in big data the large sample size has its own relative advantage in terms of exploiting heterogeneity in an advantageous way to understand the association between certain covariates (i.e. size of the organization, time, absorptive capacity of the organization, organizational compatibility) and rare outcomes such as sudden increase or decrease in market share or profitability of the organization and understanding how sustainable practices adopted by the organizations can help them to perform better than their competitors. We present the mixture model for the population as:

$$\mu 1\, p\, 1\ (y;\ \theta\, 1(x)) + \ldots\ldots\ldots\ldots\ldots\ldots + \mu m\, p\, m\ (y; \theta\, m(x)), \quad (1)$$

where $\mu j \geq 0$ represents the proportion of the jth subpopulation $p\, j$ and $(y;\ \theta\ m(x))$ is the probability distribution of the response of the jth subpopulation given the covariates x with $\theta j\ (x)$ as the parameter vector. In reality, many subpopulations rarely exist, i.e. μj is very small ($\mu j \rightarrow 0$) making it infeasible to infer the covariate-dependent parameters $\theta j(x)$ due to lack of information. However in big data due to large sample size (n), the sample size $n^*\mu j$ for the jth subpopulation can be moderately large even if μj is very small. This enables us to infer about the subpopulation parameter θj (.).

Besides the aforementioned advantages, the heterogeneity of big data may also pose significant challenges as far as statistical inference is concerned. Hence to draw an inference from mixture model as shown in equation 1 for large datasets requires sophisticated statistical and computational methods. Fan et al. (2014) argued that in case of low dimensions, standard techniques such as expectation-maximization in case of mixture model can be applied effectively. Khalili and Chen (2007) and Stadler et al. (2010) have noted that in case of high dimensions, we need to be careful while estimating parameters to avoid over fitting or noise accumulations. In our case we have determined the heterogeneity using Higgins' (2003) equation $I^2 = ((Q\text{-}df)/Q)^*100$ %, where Q represents chi-squared statistics and df represent degrees of freedom. In our case the I^2 value obtained is greater than 90%. Hence we can conclude that there exists considerable heterogeneity in our dataset. However in the past, heterogeneity in a dataset was argued as a limitation due to multiple reasons such as compromise with internal and external validity (Becker et al. 2013). However we argue that in legacy of big data, heterogeneity can be useful in exploring interesting observations that were not explored using traditional datasets. Hence we believe that a good computation algorithm needs to be designed.

5.1.2 Noise Accumulation

While dealing with BD, we need to estimate various parameters or test these parameters. These estimation errors accumulate when a decision is based on large parameters. Such a noise accumulation effect is especially severe in high dimensions and may even dominate the true signals (Fan et al. 2014). Such cases are usually handled by sparsity assumption. Hence based on the arguments offered by Fan et al. (2014) we have used sparse models and variable selections to overcome these difficulties.

Noiseless observations

Consider a linear system of equations, say $X = D^* \omega$, where D is an undetermined m*p matrix (m ≤ p) and ω R^p. D, is called the design matrix. The problem is to estimate the signal ω, subject to the constraint that it is sparse. The underlying motivation for sparse decomposition problems is that even though the observed values are high dimensional (m) space, the actual signal is organized in some lower-dimensional subspace (k<< m). The sparsity implies that only few components of ω are non-zero and rest are zero.

The sparse decomposition problem is represented as,

$$\min \omega \in R^p \ \|\omega\| 0 \text{ such that } X = D^* \omega, \qquad (2)$$

Where $\|\omega\|0 = \neq \{i: \omega i \neq o, i=1,\ldots\ldots,p\}$ is a pseudo-norm.

Noisy observations

$$\min_{\omega \in R^p} 1/2\|X - D * \omega\| + \lambda \|\omega i\|1, \qquad (3)$$

where λ is a slack variable and $\|\omega\|1$ is the sparsity-inducing term. The slack variable balances the trade-off between fitting the data perfectly and employing a sparse solution.

5.1.3 Spurious Correlation

In case of big data the large dimensionality gives rise to a problem of spurious correlation, referring to the fact that many uncorrelated random variables may have high sample correlations in high dimensions. Hence if spurious correlations were not properly taken care of, it may lead to false scientific discoveries and wrong statistical inferences as argued by Fan et al. (2014).

Consider the problem of estimating the coefficient vector β of a linear model

$$Y=X^* \beta + \in, \qquad Var(\in) = \sigma^2 Id \qquad (4)$$

Where $Y \in R^n$ represents response vector $X= [X1, X2, X3, \ldots, Xn]T \in R^{n*d}$ represents the design matrix, $\in R^n$ represents an independent random noise vector and Id is the d*d identity matrix.

Besides variable selection, spurious correlation may lead to wrong statistical inference. This can be explained by linear equation as (4).

5.1.4 Incidental Endogeneity

Incidental endogeneity is of concern in cases of high dimensional datasets. Fan and Liao (2014) argued that most research in the field of high dimensional datasets is based on the assumption that none of the regressors are correlated with the regression error, i.e. they are exogenous. However, incidental endogeneity arises easily in a large pool of regressors in a high-dimensional regression. The occurrence of incidental endogeneity may impact upon the final research conclusion.

To explain we present the regression equation as $Y= \sum \beta j^* Xj + \varepsilon$, and

$E(\varepsilon *Xj)=0$ for $j=1,2,3,4,....,d$. (5)

With a small set $S= \{j: \beta j \neq 0\}$. The exogenous assumption in equation (5) that the residual noise ε is uncorrelated with all the predictors is crucial for the validity of most existing statistical procedures, including variable selection consistency.

As we have seen, the characteristics of big data (high sample size and high dimensionality) introduce heterogeneity, noise accumulation, spurious correlation and incidental endogeneity. These characteristics of big data make traditional statistical methods invalid. Hence we attempted to check all the properties before we moved on.

5.2 Dimension Reduction and Random Projection

Golub and Van Loan (2012) argued that in the case of a high dimensionality data set, data reduction using the most popular technique (i.e. principal component analysis) is quite challenging. When projecting (n*d) data matrix D to this linear subspace that to obtain as (n*k) data matrix. This procedure is optimal among all the linear projection methods in minimizing the squared error introduced by projection (Fan et al. 2014). Conducting the eigen space decomposition on the sample covariance matrix is a computational challenge when both n and d are large. The computational complexity of PCA is

$o(d^2n + d^3)$ (Golub and Van Loan, 2012; Fan et al. 2014), which is not feasible in case of large datasets. Hence in such case "random projection (RP)" is recommended to use for data reduction. However in our case due to limited sample size we used both procedures (i.e. PCA and RP) and the final outcome was not different. Hence we have proceeded with PCA output. However in case of large data sets then RP would have been the better technique in comparison to PCA.

6. Data Analysis and Findings

In this section we will discuss psychometric properties of measuring items and test the research hypotheses.

6.1 Assessment of statistical properties

We performed tests for the assumptions of constant variance, existence of outliers, and normality of the gathered data to ensure that the data can be used for psychometric properties testing (e.g. Chen and Paulraj, 2004; Dubey et al. 2015). We used plots of residuals by predicted values, rankits plot of residuals and statistics of skewness and kurtosis (Eckstein et al. 2015; Dubey et al. 2015). To detect multivariate outliers, we used Mahalanobis distances of predicted variables (Cohen et al. 2003). The maximum absolute value of skewness is found to be less than 2 and the maximum absolute value of kurtosis is found to be less than 5, which is found to be well within acceptable limits (Curran et al. 1996). To ensure that multicollinearity was not a problem, we calculated variance inflation factors (VIF). All the VIFs were less than 1.5 and therefore considerably lower than the recommended threshold of 10.0 (Hair et al. 1998), suggesting that multicollinearity was not a problem. We used confirmatory factor analysis (CFA) to establish convergent validity and unidimensionality of factors as shown in Tables 3 and 4.

Table 3: Scales and their items (factor loadings, error, AVE)

Constructs with Cronbach Alpha value	Indicators	λi	SCR*	AVE
Leadership (X1) Alpha: 0.947	Well defined environmental policy	0.897	0.94	0.69
	Awareness about environmental policy	0.866		
	Top management support	0.798		
	Top management has approved special fund for investment in cleaner technologies	0.821		
	Top management positive attitude towards green practices	0.811		
	Senior managers motivate and support new ideas received from junior executives	0.813		
	Recognition of employees	0.813		
Regulatory Pressures (X2) Alpha: 0.885	Regional pollution control board pressurizing the firm to adopt green practices	0.89	0.91	0.71
	Government regulations provide clear guidelines in controlling pollution level	0.824		
	Pollution control board strictly monitors the pollution level of firm on a periodic basis	0.814		
	Green practices decrease incidence of penalty fee charged by pollution control	0.834		

	board			
Supplier Relationship Management (X3) Alpha: 0.960	Environmental criteria considered while selecting suppliers	0.878	0.93	0.74
	Firm considers environment collaboration with suppliers	0.843		
	Firm has technological integration with suppliers	0.816		
	Firm trains and educates suppliers in implementing ISO14001	0.878		
	Evironmental audit for suppliers done periodically	0.876		
Employee Involvement (X4) Alpha	Strategic participation	0.781	0.87	0.70
	Organizational participation	0.846		
	Task discretion	0.872		
Customer Relationship Management (X5) Alpha: 0.787	Does green practices improve customer satisfaction	0.821	0.90	0.70
	Do your firm recover end of life products from customers	0.837		
	Customers suggestion are implemented	0.812		
	Do your customers appreciate eco-friendly products	0.869		
Total Quality Management	Firm has successfully implemented Total Quality Management	0.818	0.90	0.69

(X6) Alpha: 0.715	Green practices promote product quality	0.813		
	Employee training for quality	0.868		
	Supplier quality assurance and management	0.834		
Total Productive Maintenance (X7) Alpha: 0.926	Maintenance strategy and policy deployment ownership	0.856	0.92	0.69
	Process/equipment classification, standardization and improvement	0.876		
	Process quality maintenance	0.897		
	Maintenance practices/procedures/practices	0.813		
	Standardization of materials, machines and methods (3Ms)	0.678		

Lean Manufacturing (X8) Alpha: 0.76	JIT tools, techniques and processes	0.762	0.87	0.56
	Standardized work/ standard operations	0.791		
	Cycle time/lead time/lot-size reduction	0.786		
	Cellular manufacturing/focused factory	0.716		
	Pull system	0.691		
Environmental Performance (Y1) Alpha: 0.881	Environmental technology	0.856	0.86	0.62
	Recycling efficiency	0.823		
	Eco packaging	0.875		
	Level of process management which includes pollution control, waste emissions, carbon footprint etc.	0.541		
Social Performance (Y2) Alpha: 0.781	Management commitment	0.858	0.85	0.65
	Customer satisfaction	0.798		
	Employee development	0.765		
Economic Performance (Y3) Alpha: 0.981	Environmental cost	0.73	0.84	0.64
	Supply chain cost	0.87		
	Return on Asset	0.789		

**Here SCR (Scale Composite Reliability)=* $(\sum \lambda i)2/((\sum \lambda i)2+(\sum ei))$
Where λi *= standard loadings of ith item;*
ei= 1- $((\sum \lambda i)2)$ *which represents the measurement error in ith item*
(Note: Detailed discussion on computation algorithm related to SCR and AVE is discussed by Fornell and Larcker (1981).

From Table 3, we can see that each scale possesses SCR>0.7 & AVE>0.5 which is above the threshold value suggested for each construct (Hair et al. 1998). The observed value of $\boldsymbol{\lambda}$i >0.5. The value is more than threshold value of each item that constitute a construct of framework shown in Figure 1. Therefore we can assume that convergent validity exists in our framework.

We have further derived Pearson's correlation coefficients as shown in Table 4.

Table 4: Pearson's correlation coefficients

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3
X1	0.83a										
X2	.052	0.84a									
X3	.009	.221**	0.86a								
X4	-.022	.051	.135*	0.83a							
X5	.040	.339**	.280**	.166**	0.84a						
X6	.080	.140*	.380**	.331**	.227**	0.83a					
X7	.008	.177**	.329**	.162**	.225**	.160**	0.75a				
X8	.127*	.306**	.323**	.127*	.228**	.211**	.114	0.79a			
Y1	.052	0.41	.221**	.051	.339**	.140*	.177**	.306**	0.79a		
Y2	.009	.221**	0.38	.135*	.280**	.380**	.329**	.323**	.221**	0.81a	
Y3	-.022	.051	.135*	1.000**	.166**	.331**	.162**	.127*	.051	.135*	0.80a

**Significant at $p<0.05$*
***Significant at $p<0.01$*
a The square root of the construct's AVE is provided along the diagonal

We compared the squared correlation between two latent constructs to their average variance extracted (AVE) (Fornell and Larcker, 1981). Discriminant validity exists when the squared correlation between each pair of constructs is less than the AVE for each individual construct, further establishing discriminant validity.

6.2 Goodness of Fit (GoF) of the Model

Tenenhaus et al. (2005) have proposed only one measure for GoF in PLS (Partial Least Square) based structural equation modeling (SEM). Since the seminal article by Tenenhaus et al. (2005) there is an increasing trend among researchers to use PLS-based SEM to test their theories. We have used the average R-Square and geometric mean of AVE for the endogenous constructs in the following formula:

GoF= Sqrt ((Average R-Square)* Geometric mean of AVE))

(Here Sqrt = square root and AVE= Average Variance Extracted)

Table 5: Goodness of Fit

Construct	R-Square (model1) Environmental Performance	R-Square (model2) Social Performance	R-Square (model3) Economic Performance	AVE
Leadership	0.154	0.180	0.207	0.69
Regulatory Pressure	0.404	0.276	0.361	0.71
Supplier Relationship Management	0.576	0.415	0.490	0.74
Employee Involvement	0.527	0.424	0.454	0.70
Customer relationship Management	0.473	0.364	0.356	0.70
Total Quality Management	0.287	0.107	0.196	0.69
Total Productive Maintenance	0.5	0.364	0.386	0.69
Lean Manufacturing	0.296	0.293	0.303	0.56
GoF	**0.52**	**0.45**	**0.48**	

Table 8 shows that the GoF for model 1 (i.e. when exogenous construct is environmental performance) is 0.52. As per Wetzels et al. (2009), if GoF is greater than 0.36 then the adequacy of the model validity is large. Similarly we calculated GoF for model 2 (i.e. social performance as exogenous construct) and model 3 (i.e. economic performance as exogenous construct). The GoF value for model 2 is 0.45 and model 3 is 0.48. Hence we can see from calculated values of GoF that the adequacies of the model validity are high.

7. Conclusion, Contributions and Further Research Directions

In the current paper we have attempted to revisit the role of BD on WCSM by using BD, which is characterized by volume, variety, velocity and veracity. The SNS offers an immense opportunity in terms of data gathering. However due to the authenticity of data and ethical issues, we have adopted classical approach using a SNS platform. We have generated a theoretical framework (see Figure 1) using extensive literature review of current literature and further tested our theoretical framework using gathered data. We have checked the psychometric properties of measurement items of our instrument. The CFA output suggests that our framework constructs possesses convergent validity and discriminant validity. Thus our constructs satisfy content validity and construct validity, which is unique from methodological point of view.

7.1 Academic and managerial contribution

This paper contributes to the literature of BD and WCSM (Whetten, 1989). Our study is a response to the call by BD scholars (Agarwal and Dhar, 2014; Dutta and Basu, 2015) for more studies on the opportunities enabled by BD. We stated the importance of BDA through our proposed framework, driven by the need expressed by scholars (e.g. Dubey et al, 2015; Wamba et al., 2015) to utilize BD to achieve superior performance according to the tenets of WCSM, but at the same time to consider the environmental and social consequences of these organizational actions. We extended the WCM term (Flynn et al., 1999) to include sustainable manufacturing and sustainable practices (e.g. Lovins et al., 1999; Despreisse et al., 2013), addressing the need expressed by Garetti and Taisch (2012) to process large data related to the environmental, social, and economic implications of WCM. Our research differs from recent studies (e.g. Opresnik and Taisch, 2015) in that we are not only focusing on the dimension of 'value', and we do not study servitisation; rather, we use 'volume', 'variety', 'velocity', and 'veracity'. Finally, our paper extends studies that focus on only operational strategies and techniques to achieve competitive advantage (Rusinko, 2007; Kannegiesser and Gunther, 2014; Dubey et al., 2015) by presenting the role of BDA in WCSM through an extensive literature review, through which particular factors are extracted, studied, and tested to create a framework that denotes the role of BDA within WCSM.

Our results provide useful lessons for practice in that they suggest that the role of BDA within WCSM to achieve superior economic, social, and environmental performance, by focusing on the factors extrapolated on our framework (Figure 1). Furthermore, they highlight the role of BDA as drivers of WCSM practices in the Indian and hence developing countries context. Today environmental concerns have triggered the need for sustainable practices, but at the same time aiming at achieving superior performance, as highlighted by WCSM. Managers could also use the framework we suggest as 'aide memoire' to assess the factors that are important to achieve WCSM through BDA.

7.2 Limitations and Further Research Directions

Our present study has its own limitations. First, we have attempted to collect data from SNS. The sample size may need to be increased. Second the data is gathered using a structured questionnaire. The analyses of the data would have been quite challenging if we had gathered data using different methods. Then the heterogeneity would have posited some different level of challenge. We argue the heterogeneity challenge: it would have offered us multiple opportunities to explore the microstructure with far more detail which in the present case the fine grain boundaries of the structure are not properly understood. Third, data reduction would have offered us enough opportunity to identify more enablers of WCSM. Fourth, we did not explore the role of BDA capabilities in WCSM. Looking at the best constituent of the BD capability (e.g., IT, HR) for improved firm performance should be part of future research directions. Indeed, prior studies suggested that competitive advantage is achieved through the firm's ability to deploy and use of distinctive, valuable, and inimitable resources and capabilities (Bhatt and Grover, 2005). In the present study we highlighted the role of BD on WCSM. The application of BDA can be largely used in the field of supply chain network design in terms of rationalization of warehouse footprints, reducing supply chain risk by improving prediction of unpredictable disasters, vehicle routing and improving customer service by reducing stock out and managing product life cycle. Fawcett and Waller (2014) have argued in their seminal work that there are five emerging "game changers" that can redefine the operations management field as: (1) BD and predictive analytics, (2) additive manufacturing, (3) autonomous vehicles, (4) materials science, and (5) borderless supply chains. They have also suggested four forces that impede transformation to higher levels of value co-creation: (1) supply chain security, (2) failed change management, (3) lack of trust as a governance mechanism, and (4) poor understanding of the "luxury" nature of corporate social responsibility initiatives. The use of BD can further help to address the four identified concerns. Hence we argue that future research should embrace BDA to redefine the future focus of the advanced manufacturing technology. Using BD new innovations can be made, for instance in terms of developing new materials such as biodegradable materials which cause less harm to the environment and can play significant role in improving the life of people.

Appendix 1

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	17.166	39.014	39.014	17.166	39.014	39.014
2	2.813	6.393	45.407	2.813	6.393	45.407
3	2.255	5.125	50.532	2.255	5.125	50.532
4	1.942	4.414	54.946	1.942	4.414	54.946
5	1.626	3.696	58.641	1.626	3.696	58.641
6	1.457	3.312	61.953	1.457	3.312	61.953
7	1.409	3.202	65.155	1.409	3.202	65.155
8	1.263	2.871	68.027	1.263	2.871	68.027
9	1.186	2.696	70.722	1.186	2.696	70.722
10	1.116	2.537	73.259			
11	1.030	2.341	75.600			
12	.969	2.203	77.804			
13	.872	1.983	79.786			
14	.795	1.807	81.593			
15	.774	1.760	83.353			
16	.706	1.604	84.958			
17	.626	1.423	86.381			
18	.584	1.328	87.709			
19	.562	1.276	88.985			
20	.481	1.094	90.079			
21	.444	1.010	91.089			
22	.415	.943	92.032			
23	.342	.778	92.811			
24	.319	.725	93.536			
25	.311	.707	94.243			
26	.300	.682	94.925			
27	.277	.631	95.555			
28	.251	.570	96.125			
29	.228	.518	96.643			
30	.204	.465	97.108			
31	.174	.395	97.502			
32	.154	.350	97.853			
33	.134	.304	98.157			
34	.119	.270	98.426			
35	.113	.258	98.684			

36	.107	.243	98.927			
37	.092	.210	99.137			
38	.082	.186	99.323			
39	.070	.158	99.481			
40	.060	.136	99.617			
41	.054	.123	99.740			
42	.046	.106	99.846			
43	.039	.089	99.934			
44	.029	.066	100.000			

Extraction Method: Principal Component Analysis.

References

Accenture (2013). The role of big data and analytics in the developing world. Report, available at: http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-ADP-Role-Big-Data-And-Analytics-Developing-World.pdf Accessed on 8th June 2015.

Ageron, B., Gunasekaran, A., & Spalanzani, A. (2012). Sustainable supply management: An empirical study. *International Journal of Production Economics*, *140*(1), 168-182.

Atlas, M. and R. Florida. 1998. "Green Manufacturing." In *Handbook of Technology Management,* edited by R. Dorf, CRC Press.

Azevedo, S. G., Carvalho, H., & Cruz Machado, V. (2011). The influence of green practices on supply chain performance: a case study approach. *Transportation research part E: logistics and transportation review*, *47*(6), 850-871.

Azzone, G., & Noci, G. (1998). Identifying effective PMSs for the deployment of "green" manufacturing strategies. *International Journal of Operations & Production Management*, *18*(4), 308-335.

Bai, C. & J. Sarkis. 2010. Greener Supplier Development: Analytical Evaluation Using Rough Set Theory. *Journal of Cleaner Production*, 17 (2): 255–264.

Baines, T., S. Brown, O. Benedettini & P. Ball. (2012). Examining Green Production and its Role within the Competitive Strategy of Manufacturers. *Journal of Industrial Engineering and Management*, 15 (1): 53–87.

Becker, J. M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly*, *37*(3), 665-694.

Belekoukias, I., Garza-Reyes, J. A., & Kumar, V. (2014). The impact of lean methods and tools on the operational performance of manufacturing organisations. *International Journal of Production Research*, *52*(18), 5346-5366.

Berg, B. L., Lune, H., & Lune, H. (2004). *Qualitative research methods for the social sciences* (Vol. 5). Boston, MA: Pearson.

Berkel, V. (2007).Cleaner Production and Eco-efficiency in Australian Small Firms. *International Journal of Environmental Technology and Management*, 7 (5/6): 672–693.

Bhateja, A. K., Babbar, R., Singh, S., & Sachdeva, A. (2012).Study of the Critical factor Finding's regarding evaluation of Green supply chain Performance of Indian Scenario for Manufacturing Sector. *International Journal of Computational Engineering& Management*, *15*(1), 74-80.

Bi, Z., & Cochran, D. (2014). Big Data Analytics with Applications. *Journal of Management Analytics*, 1(4), 249-265.

Bierma, T.J. and F.L. Wasterstraat. (1999). Cleaner Production from Chemical Suppliers: Understanding Shared Savings Contracts. *Journal of Cleaner Production* 7(2), 145–158.

Bou-Llusar, J. C., Escrig-Tena, A. B., Roca-Puig, V., & Beltrán-Martín, I. (2009). An empirical assessment of the EFQM excellence model: evaluation as a TQM framework relative to the MBNQA model. *Journal of Operations Management*, *27*(1), 1-22.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662-679.

Braunscheidel, M. J., & Suresh, N. C. (2009). The organizational antecedents of a firm's supply chain agility for risk mitigation and response. *Journal of Operations Management*, *27*(2), 119-140.

Brown, S., Squire, B., & Blackmon, K. (2007). The contribution of manufacturing strategy involvement and alignment to world-class manufacturing performance. *International Journal of Operations & Production Management*,*27*(3), 282-302.

Brown, B., Chui, M., & Manyika, J. (2011). Are you ready for the era of 'big data'. *McKinsey Quarterly*, 47(4), 24-35.

Bughin, J., Chui, M., & Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*, *56*(1), 75-86.

Cândido, C. J., & Santos, S. P. (2011). Is TQM more difficult to implement than other transformational strategies?. *Total Quality Management & Business Excellence*, *22*(11), 1139-1164.

Carter, C. R., & Rogers, D. S. (2008). A framework of sustainable supply chain management: moving toward new theory. *International journal of physical distribution & logistics management*, *38*(5), 360-387.

Chae, B. K. (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, 247-259.

Chen, I. J., & Paulraj, A. (2004). Towards a theory of supply chain management: the constructs and measurements. *Journal of operations management*, *22*(2), 119-150.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, *36*(4), 1165-1188.

Chien, M. K., and Shih, L. H. (2007). An empirical study of the implementation of green supply chain management practices in the electrical and electronic industry and their relation to organizational performances. *International Journal of Environmental Science and Technology*, 4(3), 383-94.

Churchill Jr, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(1),64-73.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. 3rd ed. Hillsdale, NJ: Erlbaum.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis.*Psychological methods*, *1*(1), 16-29.

Davenport, T.H., (2012). The human side of Big Data and high-performance analytics. *International Institute for Analytics*, pp.1–13.

Deif, A.M. (2011). A System Model for Green Manufacturing. *International Journal of Cleaner Production* 19 (14): 1553–1559.

Demirbag, M., Tatoglu, E., Tekinkus, M., & Zaim, S. (2006). An analysis of the relationship between TQM implementation and organizational performance: evidence from Turkish SMEs. *Journal of Manufacturing Technology Management*, *17*(6), 829-847.

Demirkan, H., & Delen, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, *55*(1), 412-421.

Despeisse, M., Ball, P.D., Evans, S., & Levers, A. (2012). Industrial Ecology at Factory Level – a Conceptual Model. *Journal of Cleaner Production* 31(3–4), 30–39.

Diaz-Elsayed, N., Jondral, A., Greinacher, S., Dornfeld, D., & Lanza, G. (2013). Assessment of lean and green strategies by simulation of manufacturing systems in discrete production environments. *CIRP Annals-Manufacturing Technology*, *62*(1), 475-478.

Dijcks, J.-P., 2013. *Oracle: Big Data for the Enterprise*. Redwood Shores, Oracle.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, *48*(2), 147-160.

Dubey, R., Gunasekaran, A. & Chakrabarty, A. (2015). World-class sustainable manufacturing: framework and a performance measurement system. *International Journal of Production Research.* **DOI:**10.1080/00207543.2015.1012603

Dües, C. M., Tan, K. H., & Lim, M. (2013). Green as the new Lean: how to use Lean practices as a catalyst to greening your supply chain. *Journal of cleaner production*, *40*, 93-100.

Dutta, D., & Bose, I. (2015). Managing a big data project: The case of Ramco Cements limited. *International Journal of Production Economics*. doi:10.1016/j.ijpe.2014.12.032

Eckstein, D., Goellner, M., Blome, C., & Henke, M. (2015). The performance impact of supply chain agility and supply chain adaptability: the moderating effect of product complexity. *International Journal of Production Research*. **DOI:**10.1080/00207543.2014.970707

Eltayeb, T., Zailani, S. and Ramayah, T. (2011). Green supply chain initiatives among certified companies in Malaysia and environmental sustainability: investigating the outcomes, *Resources, Conservation and Recycling*,55,495-506.

Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National science review*, *1*(2), 293-314.

Fan, J., & Liao, Y. (2014). Endogeneity in high dimensions. *Annals of statistics*, *42*(3), 872-917.

Farish, M. (2009). Plants that are green [Toyota's lean manufacturing]. *Engineering & Technology*, *4*(3), 68-69.

Fawcett, S. E., & Waller, M. A. (2014). Supply Chain Game Changers—Mega, Nano, and Virtual Trends—And Forces That Impede Supply Chain Design (ie, Building a Winning Team). *Journal of Business Logistics*, *35*(3), 157-164.

Flammer, C. (2013). Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors. *Academy of Management Journal*, 56(3), 758-781.

Flynn, B. B., Schroeder, R. G., & Flynn, E. J. (1999). World class manufacturing: an investigation of Hayes and Wheelwright's foundation. *Journal of operations management*, *17*(3), 249-269.

Flynn, B. B., Schroeder, R. G., Flynn, E. J., Sakakibara, S., & Bates, K. A. (1997). World-class manufacturing project: overview and selected results.*International Journal of Operations & Production Management*, *17*(7), 671-685.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of marketing research*, 18(1),39-50.

Franchetti, M., Bedal, K., Ulloa, J., & Grodek, S. (2009). Lean and green-industrial engineering methods are natural stepping stones to green engineering. *Industrial Engineer*, *41*(9), 24.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*,*35*(2), 137-144.

Garbie, I. H. (2013). DFSME: design for sustainable manufacturing enterprises (an economic viewpoint). *International Journal of Production Research*, *51*(2), 479-503.

Garbie, I. H. (2014). An analytical technique to model and assess sustainable development index in manufacturing enterprises. *International Journal of Production Research*, *52*(16), 4876-4915.

Garetti, M., & Taisch, M. (2012). Sustainable manufacturing: trends and research challenges. *Production Planning & Control*, *23*(2-3), 83-104.

Gavronski, I., Paiva, E. L., Teixeira, R., and de Andrade, M. C. F. (2013). ISO 14001 certified plants in Brazil–taxonomy and practices. *Journal of Cleaner Production*, *39*, 32-41.

Geffen,D., Straub,D.W., & Bourdreau, M.C.(2000).Structural Equation Modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*,4(7),1-79.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research,25*(2),186-192.

Gobble, M. M. (2013). Big Data: The Next Big Thing in Innovation. *Research Technology Management*. 56(1): 64-66.

Golub, G. H., & Van Loan, C. F. (2012). *Matrix computations* (Vol. 3). JHU Press

Gong, Q., Yang, Y., & Wang, S. (2014). Information and decision-making delays in MRP, KANBAN, and CONWIP. *International Journal of Production Economics*, *156*, 208-213.

Gunasekaran, A. and A.Spalanzani.2012.“Sustainable of manufacturing services: Investigation for research and applications”, *International Journal of Production Economics*, 140(1):35-47.

Gunasekaran, A., Irani, Z., and Papadopoulos, T. (2013). Modelling and analysis of sustainable operations management: certain investigations for research and applications. *Journal of the Operational Research Society*.

Gunn, T. G. (1987). *Manufacturing for competitive advantage: becoming a world class manufacturer*. Boston MA: Ballinger publishing company.

Hair, J.F., R.E. Anderson, R.L. Tatham and W.C Black. (1998).*Multivariate Data Analysis*, Upper Saddle River, NJ: Prentice-Hall.

Hajmohammad, S., Vachon, S., Klassen, R. D., & Gavronski, I. (2013). Lean management and supply management: their role in green practices and performance. *Journal of Cleaner Production*, *39*, 312-320.

Hall, R.W.(1987).Attaining Manufacturing Excellence: Just-in-Time, Total Quality, Total People Involvement, Dow Jones-Irwin, Homewood, IL

Hargittai, E. (2007). Whose space? Differences among users and non - users of social network sites. *Journal of Computer - Mediated Communication*, *13*(1), 276-297.

Hayes, R. H., & Wheelwright, S. C. (1984). *Restoring our competitive edge: competing through manufacturing* (Vol. 8). New York: Wiley.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72-80.

Hines, P., Holweg, M., & Rich, N. (2004). Learning to evolve: a review of contemporary lean thinking. *International Journal of Operations & Production Management*, *24*(10), 994-1011.

Hsu, C.W. and Hu, A.H. 2008. "Green Supply Chain Management in the Electronic Industry." *International Journal of Science and Technology*, 5 (2): 205–216.

Hsu, C.W. and Hu.A.H.2009. Applying Hazardous Substance Management to Supplier Selection Using Analytic Network Process. *Journal of Cleaner Production* 17 (2): 255–264.

International Trade Administration, (2007), How Does Commerce Define Sustainable Manufacturing? U.S. Department of Commerce. Available: http://www.trade.gov/competitiveness/sustainablemanufacturing/how_doc_defines_SM.asp Accessed on 8th June 2015.

Jabbour, C. J. C., Jabbour, A. B. L. D. S., Govindan, K., Teixeira, A. A., & Freitas, W. R. D. S. (2013). Environmental management and operational performance in automotive companies in Brazil: the role of human resource management and lean manufacturing. *Journal of Cleaner Production*, *47*, 129-140.

Jacobs, A. (2009). The pathologies of big data. *Communications of the ACM*, 52 (8), 36.

Jasiulewicz-Kaczmarek, M. (2013). Sustainability: Orientation in Maintenance Management—Theoretical Background. In *EcoProduction and Logistics* (pp. 117-134). Springer Berlin Heidelberg.

Jovane, F., Yoshikawa, H., Alting, L., Boër, C. R., Westkamper, E., Williams, D., ... & Paci, A. M. (2008). The incoming global technological and industrial revolution towards competitive sustainable manufacturing. *CIRP Annals-Manufacturing Technology*, *57*(2), 641-659.

Jayaram, J., Vickery, S., & Droge, C. (2008). Relationship building, lean strategy and firm performance: an exploratory study in the automotive supplier industry. *International Journal of Production Research*, *46*(20), 5633-5649.

Kaebernick, H., Kara, S., & Sun, M. (2003). Sustainable product development and manufacturing by considering environmental requirements. *Robotics and Computer-Integrated Manufacturing*, *19*(6), 461-468.

Kannegiesser, M., & Günther, H. O. (2014). Sustainable development of global supply chains—part 1: sustainability optimization framework. *Flexible Services and Manufacturing Journal*, *26*(1-2), 24-47.

Kaynak, H. (2003). The relationship between total quality management practices and their effects on firm performance. *Journal of operations management*, *21*(4), 405-435.

Khalili, A., & Chen, J. (2007). Variable selection in finite mixture of regression models. *Journal of the american Statistical association*, *102*(479),1025-1038.

Klassen, R. D., & McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management science*, *42*(8), 1199-1214.

Ku, C. Y., Chang, C. T., and Ho, H. P. (2010). Global supplier selection using fuzzy analytic hierarchy process and fuzzy goal programming. *Quality and Quantity*, *44*(4), 623-640.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600). ACM.

Lakhal, L., Pasin, F., & Limam, M. (2006). Quality management practices and their impact on performance. *International Journal of Quality & Reliability Management*, *23*(6), 625-646.

Law, K. M., and Gunasekaran, A. (2012). Sustainability development in high-tech manufacturing firms in Hong Kong: Motivators and readiness. *International Journal of Production Economics*, *137*(1), 116-125.

Lewis, M. A. (2000). Lean production and sustainable competitive advantage.*International Journal of Operations & Production Management*, *20*(8), 959-978.

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big Data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, 1-18.

Lomborg, S., & Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, *30*(4), 256-265.

Luthra, S., Kumar, V., Kumar, S., & Haleem, A. (2011). Barriers to implement green supply chain management in automobile industry using interpretive structural modeling technique: An Indian perspective. *Journal of Industrial Engineering and Management*, *4*(2), 231-257.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity.

McKinsey Global Institute Report (http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)(accessed on 23rd January, 2014).

Mar Fuentes-Fuentes, M., Albacete-Sáez, C. A., & Lloréns-Montes, F. J. (2004). The impact of environmental characteristics on TQM principles and organizational performance. *Omega*, *32*(6), 425-442.

Mark, A. B., & Laney, D. (2012). The Importance of' Big Data': A Definition.*Gartner, Jun*, *21*.

Marsden, P. V. (1990). Network data and measurement. *Annual review of sociology*, 16, 435-463.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, New York.

McAdam, R. (1999). Life after ISO 9000: an analysis of the impact of ISO 9000 and total quality management on small businesses in Northern Ireland. *Total Quality Management*, *10*(2), 229-241.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big Data :*The management revolution. Harvard Bus Rev*, *90*(10), 61-67.

McGahan, A. (2013). *Unlocking The Big Promise of Big Data*. Totman Management, 6(1),53-57.

Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, *10*(7), 1497-1516.

Molamohamadi, Z., & Ismail, N. (2013). Developing a New Scheme for Sustainable Manufacturing. *International Journal of Materials, Mechanics and Manufacturing*, *1*(1), 1-5.

Mudgal, R. K., Shankar, R., Talib, P., & Raj, T. (2010). Modelling the barriers of green supply chain practices: an Indian perspective. *International Journal of Logistics Systems and Management*, *7*(1), 81-107.

Murovec, N., R.S. Erker, and I. Prodan, 2012. “Determinants of Environmental Investments: Testing the Structural Model.” *Journal of Cleaner Production* 37: 265–277.

Ng, I., Scharf, K., Pogrebna, G., & Maull, R. (2015). Contextual variety, Internet-of-Things and the choice of tailoring over platform: Mass customisation strategy in supply chain management. *International Journal of Production Economics*, *159*, 76-87.

Noci, G. (1997). Designing ‘green’vendor rating systems for the assessment of a supplier's environmental performance. *European Journal of Purchasing & Supply Management*, *3*(2), 103-114.

Opresnk, D., & Taisch, M. (2015). The value of Big Data in servitization. *International Journal of Production Economics*, 165, 174–184.

Pauli, G. 1997. “Zero Emissions: The Ultimate Goal of Cleaner Production.” *Journal of Cleaner Production* 5 (1/2): 109–113.

Pereira-Moliner, J., E. Claver-Cortes, J. Molina-Azorin and J. Tari. (2012). Quality Management, Environmental Management and Firm Performance: Direct and Mediating Effects in the Hotel Industry. *Journal of Cleaner Production* 37: 82–92.

Pochampally, K. K., Gupta, S. M., & Govindan, K. (2009). Metrics for performance measurement of a reverse/closed-loop supply chain. *International Journal of Business Performance and Supply Chain Modelling*, *1*(1), 8-32.

Prabhu, V. V., Jeon, H. W., & Taisch, M. (2012, August). Modeling green factory physics—An analytical approach. In *Automation Science and Engineering (CASE), 2012 IEEE International Conference on* (pp. 46-51). IEEE.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1*(1), 51-59.

Pusavec, F., Krajnik, P., & Kopac, J. (2010). Transitioning to sustainable production–Part I: application on machining technologies. *Journal of Cleaner Production*, *18*(2), 174-184.

Rao,P. and Holt,D.(2005)."Do green supply chains lead to competitiveness and economic performance?”, *International Journal of Operations and Production Management*, 25(9), 898 – 916.

Roth, A. V., Giffi, C. A., & Seal, G. M. (1992). Operating strategies for the 1990s: elements comprising world-class manufacturing. *Voss, C.(éd.), Manufacturing Strategy. Process and Content, Chapman & Hall, London*, 133-165.

Rusinko, C. A. (2007). Green Manufacturing: An evaluation of environmentally sustainable manufacturing practices and their impact on competitive outcomes. *IEEE Transactions on Engineering Management* 54(3), 445-454

Sadikoglu, E., & Zehir, C. (2010). Investigating the effects of innovation and employee performance on the relationship between total quality management practices and firm performance: An empirical study of Turkish firms.*International Journal of Production Economics*, *127*(1), 13-26.

Sánchez, A. M., & Pérez, M. P. (2001). Lean indicators and manufacturing strategies. *International Journal of Operations & Production Management*,*21*(11), 1433-1452.

Sarkis, J., Gonzalez-Torre, P., & Adenso-Diaz, B. (2010). Stakeholder pressure and the adoption of environmental practices: The mediating effect of training.*Journal of Operations Management*, *28*(2), 163-176.

Sarkis, J., Zhu, Q., and Lai, K., 2011. An organizational theoretic review of green supply chain management literature. *International Journal of Production Economics* 130(1), 1-15.

Schoenherr, T., & 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.

Schonberger, R.J (1986). World Class Manufacturing: The Lessons of Simplicity Applied, Free Press, New York.

Schroeder, R. G., & Flynn, B. B. (Eds.).(2002). *High performance manufacturing: Global perspectives*. Wiley,NY.

Seman, N. A. A., Zakuan, N., Jusoh, A., Arif, M. S. M., & Saman, M. Z. M. (2012). Green Supply Chain Management: A review and research direction. *International Journal of Managing Value & Supply Chains*, *3*(1),1-18.

Seuring, S., & Müller, M. (2008). Core issues in sustainable supply chain management–a Delphi study. *Business Strategy and the Environment*, *17*(8), 455-466.

Sharma, M., & Kodali, R. (2008). Development of a framework for manufacturing excellence. *Measuring Business Excellence*, *12*(4), 50-66.

Siaminwe, L., K. Chinsembu and M. Syakalima.(2005).“Policy and Operational Constraints for the Implementation of Cleaner Production.” *Journal of Cleaner Production* 13: 1037–1047.

Sila, I. (2007). Examining the effects of contextual factors on TQM and performance through the lens of organizational theories: an empirical study.*Journal of Operations management*, *25*(1), 83-109.

Sila, I., & Ebrahimpour, M. (2005). Critical linkages among TQM factors and business results. *International journal of operations & production management*,*25*(11), 1123-1155.

Singh, A., B. Singh and A.K. Dhingra.(2012). Drivers and Barriers of Green Manufacturing Practices: A Survey of Indian Industries. *International Journal of Engineering Sciences* 1 (1): 5–19.

Städler, N., Bühlmann, P., & Van De Geer, S. (2010). ℓ 1-penalization for mixture regression models. *Test*, *19*(2), 209-256.

Steudel, H. J., & Desruelle, P. (1992). *Manufacturing in the'90s: How to Become a Mean, Lean World-Class Competitor*. Van Nostrand Reinhold Company.

Stone, L.J. 2006. Limitations of Cleaner Production Programmes as Organizational Change Agents. II Leadership, Support, Communication, Involvement and Programme Design. *Journal of Cleaner Production* 14: 15–30.

Sun, E. W., Chen, Y. T., & Yu, M. T. (2015). Generalized Optimal Wavelet Decomposing Algorithm for Big Financial Data. *International Journal of Production Economics*. doi:10.1016/j.ijpe.2014.12.033.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational statistics & data analysis*, *48*(1), 159-205.

Testa, F. and F. Iraldo. (2010). Shadows and Lights of GSCM (Green Supply Chain Management): Determinants and Effects of these Practices Based on a Multinational Study. *Journal of Cleaner Production* 18 (10/11): 953–962.

Tsoulfas, G.T. and C.P. Pappis. (2006). Environmental Principles Applicable To Supply Chains Design and Operation. *Journal of Cleaner Production* 14 (1): 1593–1602.

Tufekci, Z. (2008). Grooming, gossip, Facebook and MySpace: What can we learn about these sites from those who won't assimilate?. *Information, Communication & Society*, *11*(4), 544-564.

van Hoof, B., & Lyon, T. P. (2013). Cleaner production in small firms taking part in Mexico's Sustainable Supplier Program. *Journal of Cleaner Production*, *41*, 270-282.

Vachon, S. and R.D. Klassen. 2006. Green Project Partnership in the Supply Chain: The Case of the Package Printing Industry. *Journal of Cleaner Production* 14 (6/7): 661–671.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77-84.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*. (DOI:10.1016/j.ijpe.2014.12.031).

Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration, *MIS quarterly*, 33(1), 177-195.

Whetten, D. A. (1989). What constitutes a theoretical contribution?. *Academy of management review*, *14*(4), 490-495.

White, M., (2012). Digital workplaces: vision and reality. *Business Information Review*, 29(4), 205–214.

Whitelock, V. G. (2012). Alignment between green supply chain management strategy and business strategy, *International Journal of Procurement Management*, *5*(4), 430-451.

Worley, J. M., & Doolen, T. L. (2006). The role of communication and management support in a lean manufacturing implementation. *Management Decision*, *44*(2), 228-245.

Zairi, M., & Peters, J. (2002). The impact of social responsibility on business performance. *Managerial Auditing Journal*, *17*(4), 174-178.

Zhang, Y., Chen, M., Mao, S., Hu, L., & Leung, V. C. (2014). CAP: community activity prediction based on big data analysis. *Network, IEEE*, *28*(4), 52-57.

Zhong, R. Y., Huang, G. Q., Lan, S., Dai, Q. Y., Chen, X., & Zhang, T. (2015). A big data approach for logistics trajectory discovery from RFID-enabled production data. *International Journal of Production Economics*, *165*, 260-272.

Zhu, Q., J. Sarkis and Y. Geng. (2005).Green Supply Chain Management in China: Pressure, Practices and Performance. *International Journal of Operations and Production Management* 25 (5): 449–468.