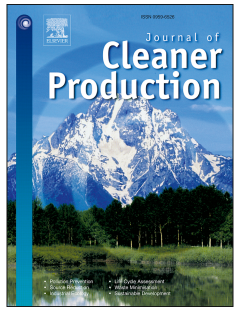
A big data driven analytical framework for energy-intensive manufacturing industries

Yingfeng Zhang, Shuaiyin Ma, Haidong Yang, Jingxiang Lv and Yang Liu

Accepted Manuscript



A big data driven analytical framework for energy-intensive manufacturing industries

Yingfeng Zhang, Shuaiyin Ma, Haidong Yang, Jingxiang Lv, Yang Liu

PII: S0959-6526(18)31820-1

DOI: 10.1016/j.jclepro.2018.06.170

Reference: JCLP 13316

To appear in: *Journal of Cleaner Production*

Received Date: 29 September 2017

Revised Date: 4 June 2018

Accepted Date: 15 June 2018

Highlights

A big data driven analytical framework for energy-intensive industries is proposed.

Useful information are mined by integrating big data and energy consumption analysis.

Energy-efficient decisions can be made based on the proposed framework.

A big data driven analytical framework for energy-intensive manufacturing industries

Yingfeng Zhang [a,b], Shuaiyin Ma [a], Haidong Yang [c], Jingxiang Lv [a,]*, Yang Liu [d,e,]*

[a] Key Laboratory of Contemporary Design and Integrated Manufacturing Technology, Ministry of Education, Northwestern Polytechnical University, Shaanxi, P. R. China, 710072

[b] Research & Development Institute in Shenzhen, Northwestern Polytechnical University

[c] Key Laboratory of Computer Integrated Manufacturing System, Guangdong University of Technology, Guangzhou, P.R. China, 510006

[d] Division of Environmental Technology and Management, Department of Management and Engineering, Linköping University, SE-581 83 Linköping, Sweden

[e] Department of Production, University of Vaasa, PL 700,65101 Vaasa, Finland

* Corresponding Authors: lvjx@nwpu.edu.cn (J. Lv), yang.liu@liu.se (Y. Liu)

Abstract: Energy-intensive industries account for almost 51% of energy consumption in China. A continuous improvement in energy efficiency is important for energy-intensive industries. Cleaner production has proven itself as an effective way to improve energy efficiency and reduce energy consumption. However, there is a lack of manufacturing data due to the difficult implementation of sensors in harsh production environment, such as high temperature, high pressure, high acid, high alkali, and smoky environment which hinders the implementation of the cleaner production strategy. Thanks to the rapid development of the Internet of Things, many data can be sensed and collected in the manufacturing processes. In this paper, a big data driven analytical framework is proposed to reduce the energy consumption and emission for energy-intensive manufacturing industries. Then, two key technologies of the proposed framework, namely energy big data acquisition and energy big data mining, are utilized to implement energy big data analytics. Finally, an application scenario of ball mills in a pulp workshop of a partner company is presented to demonstrate the proposed framework. The results show that the energy consumption and energy costs are reduced by 3% and 4% respectively. These improvements can promote the implementation of cleaner production strategy and contribute to the sustainable development of energy-intensive manufacturing industries.

Keywords: Energy-intensive manufacturing industries, Big data analytics, Cleaner production, Data mining

1. Introduction

Under the pressure of limited natural resources and increasingly severe environmental problems, energy saving and emission reduction are two important goals for manufacturing industries, especially for energy-intensive industries (EIIs) (Liu and Wang, 2017). EIIs are of great importance for the national economic development since they produce raw materials, e.g. glass (Lechtenböhmer et al., 2016), cement (Chan et al., 2014), ceramics (Fan et al., 2017), steel (Porzio et al., 2013), pulp and paper (Thollander and Ottosson, 2010), and nonferrous metals (Lin and Tan, 2016), but they have a significant impact on resource consumption and environmental pollution. In China, the six major EIIs include Processing of Petroleum, Coking and Processing of Nuclear Fuel (PPCPNF), Manufacture of Raw Chemical Materials and Chemical Products (MRCMCP), Manufacture of Non-metallic Mineral Products (MNMP), Smelting and Pressing of Non-Ferrous Metals (SPNM), Smelting and Pressing of Ferrous Metals (SPFM), as well as Production and Supply of Electric Power and Heat power (PSEPHP) (China Economic and Social Development Statics Bulletin, 2010, Wang et al. 2015, Li et al., 2014). China is the world's largest energy consumer, accounting for 23% of global energy consumption and contributing 27% to global energy demand growth in 2016 (BP Statistical Review, 2017). Fig. 1 depicts the latest energy consumption of EIIs and its proportions to the whole industrial and national levels in the period 2000-2015 in China. In 2000, the total energy consumption in China's EIIs was 691.68 million tons of standard coal equivalent (SCE), accounting for 66.65%

of total industrial energy consumption, and for 47.07% of China's total energy consumption in that year. In 2015, this part of energy has increased to 2714.69 million tons of SCE, accounting for 74.41% and 50.57% of energy consumed in industries and at national level, respectively. In the past 16 years, the total energy consumption of China's EIIs has increased 3.14 times, and its proportion in the whole industries and the national energy consumption has also increased.

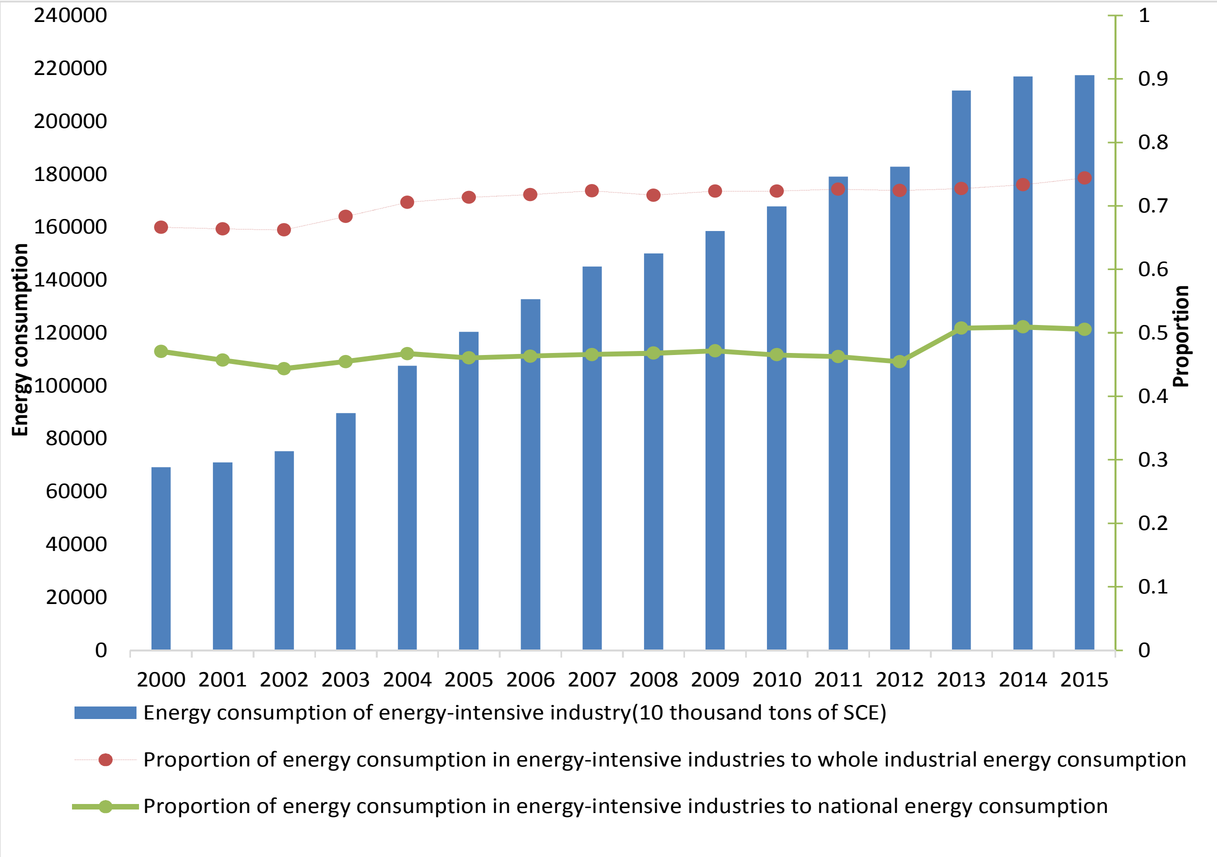


Fig. 1. Energy consumption of energy-intensive industries and its proportion to the whole industrial and national levels in China from 2000 to 2015 (National Bureau of Statistics of China, 2016).

Energy-intensive manufacturing industries (EIMIs) (Napp et al., 2014), which are EIIs in manufacturing applications, take advantage of large-scale facilities and equipment in production and have a higher energy consumption than any other industries (Song and Oh, 2015). In EIMIs, the production chain consists of continuous flow and discrete flow manufacturing processes. The interactions between the two types of processes increase the complexity of modeling and evaluating the energy consumption performance (Li et al., 2017). A deep insight into the energy consumption patterns from an individual process to the entire production chain is a prerequisite for improving the energy efficiency. However, the energy data are collected with difficulty, especially in harsh production environment in EIMIs. With the development of new technologies such as Internet of Things (IoT), the manufacturing industries could utilize the advanced information technologies such as radio frequency identification (RFID), smart sensors, and smart meters to collect the energy-consumption-related data for energy-saving and emission-reduction of products (Tao et al., 2014).

In EIMIs, manufacturing process generates massive amounts of energy data from process equipment, production process and operation management at unprecedented speed. The data is a mixture of structured (e.g., energy consumption data including spatial, time, and energy dimension), semi-structured (e.g., data exchanged between smart energy management platforms), and unstructured data (e.g. email notifications about energy use, interactions of consumers on social media about their energy use). Such energy data are characterized by high volume, high velocity, high variety and high value, which belong to a typical family of big data (Zhou et al., 2016, Zhou and Yang, 2016, Koseleva and Ropaite, 2017).

Big data refers to a collection of data sets that are too large and complex to efficiently manage and process using the traditional technologies and tools (Jacobs, 2009, Zhong et al., 2015, Zhong et al., 2017). In order to take a deep dig into the implementation of big data in manufacturing, a big data analytical architecture is proposed for cleaner production (CP) of complex products (Y. Zhang et al., 2017c). Under the architecture, manufacturers can use the advanced analytical tool to optimize the factors that are proved to have the greatest effect on CP. CP has proven itself as an effective way of obtaining the improved material utilization, reduced energy consumption and lower emission levels (Kjaerheim, 2005). The manufacturing industries can achieve the expected objective and obtain obvious progress in energy conservation and emission reduction through the application of CP (Huang et al., 2013). While manufacturing industries are struggling to improve their sustainable competitive advantage (Liu, 2013, Liu and Liang, 2015), CP has become an effective strategy which is resulting in the development of enterprise informationalization.

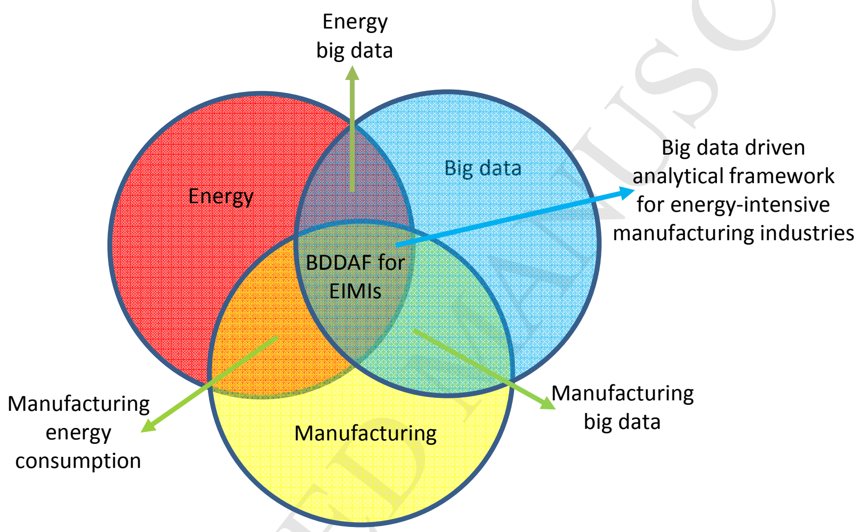


Fig. 2. Interdisciplinary research areas of energy, big data, and manufacturing

The large amounts of the energy consumption data available and the advanced techniques of big data analytics have combined to trigger the formation of a new interdisciplinary research area, namely the energy big data. The deepened research and development of energy big data analytics and its applications have brought new opportunities for understanding energy consumption in EIMIs. Energy, manufacturing, and big data intersect one another, which are no longer independent disciplines, thus forming some new crossed research areas. Fig. 2 shows the intersection of energy, manufacturing, and big data, as well as the positioning of interdisciplinary research areas, including energy big data, manufacturing energy consumption, and manufacturing big data. Currently, there is almost little scientific research on the combination of energy big data and manufacturing. Energy big data analytics in manufacturing is a new interdisciplinary research area of energy, big data, and manufacturing. Therefore, a big data driven analytical framework (BDDAF) for EIMIs is proposed in this study, so that we could provide a theoretical and practical research direction in the academic and industrial field. Considering the difficulty of the acquisition and mining of energy data mentioned above, the following research questions are of our particular interest.

1. How to establish a BDDAF for EIMIs with an integrated and systemic approach for energy conservation and emission reduction?

2. How to establish an overall energy big data perception and acquisition framework to sense the multi-source and heterogeneous energy big data, especially in harsh production environment, such as high temperature, high pressure, high acid, high alkali, and smoky environment?

3. How to discover hidden knowledge from energy big data to avoid unnecessary wastes and to overcome the shortage of energy-efficient knowledge during the implementation of CP strategy?

These research questions are addressed in the rest of the paper, which is structured as follows. A literature review is conducted in Section 2. Then the overall architecture of the BDDAF for EIMIs is proposed in Section 3, followed by the development of the key technologies related to energy big data analytics in Section 4. In section 5, an application scenario of a partner company is used to illustrate the implementation of the proposed framework. Discussions are presented in Section 6. Finally, the managerial implications and conclusions are given in Section 7 and Section 8.

2. Literature review

This section reviews related research which is categorized into two dimensions: (1) measurement of energy consumption in production, and (2) energy big data in manufacturing. The knowledge gaps are identified and summarized in the end of the section.

2.1 Measurement of energy consumption in production

The advantages of IoT technologies are becoming increasingly prominent, as they have been widely used to measure energy consumption during the whole production process. This section briefly reviews this topic of using IoT technologies to the measurement of real-time energy data. Furthermore, soft sensor approaches are also introduced subsequently for the estimation of difficult-to-measure process variables.

In the manufacturing industry, IoT technologies are enhancing the monitoring of production processes almost in real-time. An area where IoT technologies (e.g. smart meters (O'Driscoll and O'Donnell, 2013) and sensors (Bunse and Vodicka, 2010)) play a major role is in the monitoring of energy consumption (Haller et al., 2009). To be specific, the smart meters such as the electricity, gas, and water meters collect the data of electricity, gas, and water, and the sensor technologies mainly capture the energy consumption data through the parameters of temperature, pressure, etc. In reality, the energy management faces challenges in manufacturing due to the complexity that arises from the variety of energy use across the thousands of processes. So Kara et al. (2011) define three levels of energy metering in factories, which are the factory level metering, the production line-level metering, and the metering at the machine level. Based on the practices in production management, Shrouf and Miragliotta (2015) have developed a framework for the IoT-based energy management to support the integration of gathered energy into a company's production management. Moreover, an architecture for the real-time energy information capturing of the internet of manufacturing things (IoMT) has been developed to support better-informed workshop decisions (Zhang et al., 2015). Also, Tao et al. (2016) have indicated that IoT can be and is being applied in the energy management of products. Consequently, the real-time status of resources and the data of energy consumption from the manufacturing process, in theory, can be collected to improve the energy-efficient decision-making (W. Wang et al., 2017).

However, in modern industrial processes, some important variables are difficult to be measured online due to the limitations of process technology or the measurement technologies (Yan et al., 2004). Fortunately, soft sensors are employed to solve such problems. Comparing with the traditional hardware sensors, a soft sensor is a combinatorial technology of a mathematical model, data processing, and software techniques. The core of the soft sensor is the soft sensor model of a process that generates a virtual measurement to replace a real sensor measurement. There are two main families of soft sensors, i.e. the physical model based ones and data-driven ones. The physical model can be used to estimate these key measure parameters by the chemical and physical principles underlying the process between the key measure parameters and the easy-to-measure parameters (Kadlec et al., 2009). Nevertheless, a physical model is often not available because of the complexity of the machining mechanism and the large computational time requirement. As a result, the data-driven model is another way to develop the soft sensor (Shang et al., 2014). A data-driven model is a black-box model, based only on measurements in an industrial process. In the modeling procedures, the relationship between input and output of the plant can be emphasized while the knowledge of some sophisticated processes is ignored. With the development of advanced analytical tools, a wide variety of artificial intelligence and machine learning techniques have provided some powerful modeling tools for the data-driven models (Yan et al., 2017). Consequently, these important process variables are estimated by the soft sensor approaches. Nevertheless, most of the estimated variables are only used to properly control the product quality (Jian et al., 2017). The measurement of the energy consumption in extreme production conditions is seldom investigated.

2.2 Energy big data in manufacturing

Big data in the energy sector, i.e. the energy big data, also have the “4V” characteristics, namely volume, velocity, variety, and value (Zhou and Yang, 2016). This section briefly reviews the methods of energy consumption analysis and energy big data analytics in manufacturing.

In order to improve our understanding of the industrial energy use, efficiency, and pollution, Polenske and McMichael (2002) have illustrated how the input-output process model could be used to determine the economic and specific energy requirements. However, the input-output process model is only used in the cokemaking industry. Liu et al. (2006a, 2006b) have presented an in-depth quantitative analysis of energy intensity and product ratio (e-p analysis) in process industries. The e-p analysis is mainly considered of energy consumption from the perspective of material flows. From the perspective of energy flows, Cai (2009) has adopted a quantitative c-g analysis to study the energy consumption of process industry. Then it is proposed that enhancing the conversion efficiency of some energy systems and reinforcing the recycling of the wasted heat resources are the main directions of energy-saving in the future. With the development of IoT technologies, the real-time data related to energy consumption can be captured and collected. A new method based on IoT has been proposed for energy-saving and emission-reduction (Tao et al., 2014) and an IoT-based green scheduling method has been proposed for manufacturing enterprises to improve the energy efficiency and the production efficiency in the manufacturing process (Y. Zhang et al., 2017d). Nevertheless, the analysis method based on IoT are not intelligent enough to meet the challenges proposed by the smart manufacturing mode. Zhang et al. (2017a) have proposed a framework of self-organizing and self-adaptive intelligent shopfloor based on the agent technology and cyber-physical system to allocate energy and resources timely. In detail, an augmented Lagrangian coordination (ALC) method has been proposed to optimize the allocation of energy and resource for manufacturing tasks (G. Zhang et al., 2017). In order to enhance the efficiency of material handling, a cyber-physical system based smart control model (Zhang et al., 2018b) and a framework depicting the mechanism and methodology of smart production-logistics systems (Zhang et al., 2018a) have been designed to reduce the consumption of energy and time. Whereas, the analysis models above are all researched theoretically. In practice, Lv et al. (2016, 2018) have investigated the energy characteristics and the power models of the computer controlled machine tools through experimental studies. The results showed that the energy-saving potential of machining process was tremendous.

Manufacturing carries a huge number of energy data, which face challenges with the traditional analyzing methods of the energy consumption. Big data analytics are proposed to address the challenges in the industrial area (Auschitzky et al., 2014). Data mining is the most important research in big data analytics and has been widely used in the industrial area. For example, Zaki (2000) has presented a survey on large-scale parallel and distributed data mining algorithms and systems. From the aspect of continuous and incremental data mining, Fong et al. (2003) have introduced a frame metadata model to facilitate the continuous association rules generation in data mining and Lin et al. (2009) have proposed the maintenance algorithm for incremental mining based on the concept of pre-large itemsets. To mine high-utility itemsets, a binary particle swarm optimization (PSO) approach has been proposed (Lin et al., 2017). Further, Fournier-Viger et al. (2017) have surveyed recent studies on sequential pattern mining and its application. Data mining can effectively promote the implementation of CP, as well as the development of sustainable production and consumption.

CP in the era of big data will increasingly depend on the support of big data analytics (Song et al., 2017).

To manage large-scale energy data, Lee et al. (2014) have developed a prototype of a big data management system for the storing, indexing, and searching of huge-scale energy usage data. In the prototype system, an energy consumption prediction model is proposed based on penalized linear regression-based map/reduce algorithms. Moreover, Diamantoulakis et al. (2015) have summarized the state-of-the-art in the exploitation of big data tools for dynamic energy management in smart grid platforms. Further, Zhou et al. (2016) have presented a comprehensive study of big data driven smart energy management and first discuss the 4V characteristics of energy big data. With that, Zhou and Yang (2016) have provided a new way to analyze and understand the energy consumption behavior through energy big data analytics. Furthermore, a big data analytics architecture for the maintenance process of complex products has been proposed to make better CP decision (Y. Zhang et al., 2017c). A qualitative case analysis has been conducted to verify the presented architecture. The results showed that the proposed architecture benefited customers, manufacturers, and environment.

A comprehensive investigation of big data challenges for enterprise application performance management is in Rabl et al. (2012). The results of this work show that big data applications in industries can be increased. In order to utilize big data mining and advanced analytics to make manufacturing decisions more rational, Auschitzky et al. (2014) have introduced an in-depth analysis of the issues. However, there is little research on the analysis of energy consumption based on big data analytics for EIMIs. Therefore, it is necessary to establish a systemic and theoretical framework, which is combined with the methods of big data analytics and the methods of traditional energy consumption analysis to analyze and solve the problem of the waste of energy consumption for EIMIs (Z. Wang et al., 2017).

Table 1

Classification and comparison of related studies.

Aspect	Disciplinary			Subject and relate studies	Gaps	Challenges and coverage of this paper
	Energy	Big data	manufacturing			
Measurement of energy consumption in production			√	Soft sensing modeling (Yan et al., 2004; Jian et al., 2017); physical sensing model (Kadlec et al., 2009); data-driven soft sensing modeling (Shang et al., 2014; Yan et al., 2017)	Many studies mainly focused on measurement of energy consumption in manufacturing industries and soft sensor in process industries. The measurement of energy consumption is seldom investigated in extreme production conditions.	Energy management faces challenges due to the complexity that arises from the variety of energy use across thousands of processes, each one having their unique energy consumption characteristics. How to sense multi-source and heterogeneous energy big data, especially in harsh production environment for EIMIs.
	√		√	Smart meters (O'Driscoll and O'Donnell, 2013); sensors (Bunse and Vodicka, 2010); three levels of energy metering (Kara et al., 2011); energy management based on IoT (Haller et al., 2009; Shrouf and Miragliotta, 2015; Tao et al., 2016; W. Wang et al., 2017); IoMT (Zhang et al., 2015)		

Energy big data in manufacturing		√		Parallel and distributed data mining (Zaki, 2000); continuous and incremental data mining (Fong et al., 2003; Lin et al., 2009); PSO approach (Lin et al., 2017); sequential pattern mining (Fournier-Viger et al., 2017)	Most data mining applications only focused on the methods of analysis of energy consumption or the methods of big data mining. Little effort has been devoted to the combination methods of energy consumption analysis and big data mining.	Energy data mining faces the challenges of the complexity of modeling and evaluating the energy consumption performance with the increasing interactions between the continuous flow and discrete flow of energy-intensive manufacturing processes. How to excavate hidden knowledge from energy big data and obtain the energy-efficient decision-making during the implementation of CP strategy.
	√	√		Energy big data management system (Lee et al., 2014); big data tools for dynamic energy management (Diamantoulakis et al., 2015); energy big data analytics (Zhou and Yang, 2016; Zhou et al., 2016)		
	√		√	Input-output process model (Polenske and McMichael, 2002); e-p analysis (Liu et al. 2006a, 2006b); c-g analysis (Cai J., 2009); green schedule (Y. Zhang et al., 2017d); self-organizing and self-adaptive model (Y. Zhang et al., 2017a); ALC method (G. Zhang et al., 2017); smart control model (Zhang et al., 2018b); smart production-logistics systems (Zhang et al., 2018a); experimental study (Lv et al., 2016, 2018)		
		√	√	Big data can improve manufacturing (Auschitzky et al., 2014); big data for enterprise management (Rabl et al., 2012);		
	√	√	√	Big data for cleaner production (Song et al., 2017; Zhang et al., 2017c); call for papers (Z. Wang et al., 2017)		

2.3 Knowledge gaps

From this review, although the significant process has been made in the two research dimensions mentioned above, as shown in Table 1, there are still some gaps that need to be filled in.

- In terms of measurement of energy consumption in production, a large number of studies mainly focused on the measurement of energy consumption in manufacturing industries, other than EIMIs. The soft sensor approaches are only used to properly control the product quality in processing industries. In EIMIs, the measurement of energy consumption in harsh production environment is seldom investigated.
- In respect of energy big data in manufacturing, most applications of data mining only focus on the methods of the analysis of energy consumption or the methods of big data mining. Little effort has been devoted to the integrated methods of the energy consumption analysis through big data mining.

3. A big data driven analytical framework for energy-intensive manufacturing industries

With the increasing energy consumption data generated, it has become a big challenge for the traditional architecture and infrastructures to process large amounts of data within an acceptable time and resources. For this reason, big data analytics has become a key factor for companies to reveal the hidden information and to achieve competitive advantages (Chong and Shi, 2015). Big data analytics consists of several major steps. Data collection, transmission, storage, cleaning, preparation, integration, and feature selection are important procedures in the preparation phase. Then, data mining is the key step and the core content of big data analytics. Afterwards, the information or knowledge extracted from big data should be represented, visualized and applied, thus supporting the decision making and control (Zhou et al., 2016).

Based on the characteristics of the energy consumption data mentioned above and the typical infrastructure (data acquisition, storage, preprocessing, mining, decision-making, application etc.) of big data analytics, an overall BDDAF of EIMIs is designed as seen in Fig. 3. Under this framework, real-time and non-real time data of the whole energy consumption status is dynamically monitored and captured. By using the network and communication techniques such as Internet, the captured data can be transmitted to and stored in the enterprise databases. Meanwhile, data preprocessing is conducted to provide available and reliable data support for further data mining and decision-making. Finally, the mined results will provide valuable knowledge and information to implement optimal control and decision of energy conservation and emission reduction. In the left side of Fig. 3, a feedback mechanism is designed to provide the services of knowledge and information interaction for workshop in a timely fashion.

The proposed framework consists of four components from the bottom to the top, namely energy big data perception and acquisition, energy data big storage and preprocessing, energy big data mining and energy-efficient decision-making, and application services of energy big data.

3.1. Energy big data perception and acquisition

As shown in the bottom of Fig. 3, the IoT devices (e.g. smart meters, smart sensors, RFID reader, and RFID tags) are configured in the distributed and dynamic manufacturing environment to capture the multi-source and heterogeneous energy data. Meanwhile, the soft sensor is used to measure the energy data in extreme conditions, such as high temperature, high pressure, high acid, and alkali, etc. Then the collected data is transferred to enterprise databases via standard communication protocols, such as Internet RS-485/323, Modbus.

3.2. Energy big data storage and preprocessing

As mentioned above, the energy big data includes real-time and non-real-time data. A large amount of structured, semi-structured and unstructured data also contains in it. The data sets are too large and complex to efficiently store and process using traditional technologies and tools of data storage and processing. Therefore, Not only Structured Query Language (NoSQL) (Cattell, 2011) will be used to store the large-scale and disordered datasets. Storm real-time computing framework is used to process the energy data which need a high real-time processing ability (Yang et al., 2013). Meanwhile, Hadoop computing framework is used to process the non-real-time energy data (Shvachko et al., 2010). The methods of data preprocessing such as cleaning, integration, reduction, and transformation are involved in this framework.

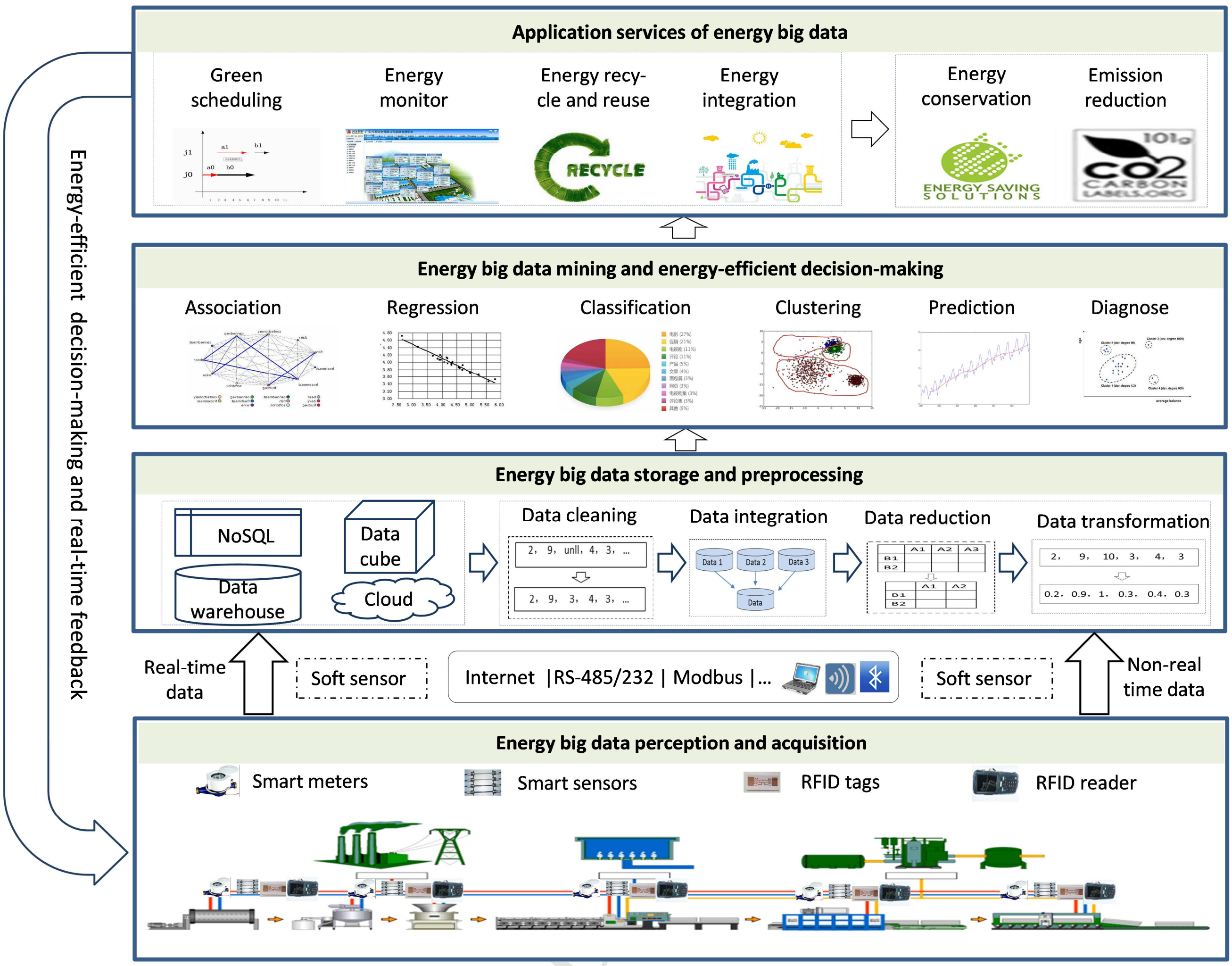


Fig. 3. A big data driven analytical framework for energy-intensive manufacturing industries (based on the example of ceramics industry).

3.3. Energy big data mining and energy-efficient decision-making

Due to the 4V (volume, velocity, variety, and value) characteristic of energy big data, it is difficult to analyze it by using the traditional methods such as e-p analysis and c-g analysis. In this context, data mining (e.g. clustering, association, classification) (Wu et al., 2014) is considered as a powerful technology that promises to discover hidden knowledge from the enormous energy data sets. By combining the methods of energy consumption analysis and the approaches of big data mining, valuable information and knowledge can be discovered from these large energy data sets. Based on the mined results, better energy-efficient decision-makings for application services will be provided to enterprise managers.

3.4. Application services of energy big data

As shown in the top of Fig. 3, application services are used to provide important real-time and non-real-time applications based on the mined information and knowledge. In order to promote the implementation of CP strategy, as well as the development of sustainable production and consumption, several types of application services (e.g. green scheduling, energy monitor, energy integration, energy recycle and reuse) are designed in this framework for the further targets of energy conservation and emission reduction. The proposed framework designs a closed-loop control process from application to workshop, which provides a real-time guidance and adjustment for the energy-efficient production.

4. Key technologies of energy big data analytics

4.1 Energy big data perception and acquisition

Table 2

Configuration information of IoT devices in manufacturing factory.

IoT devices	Manufacturing resources	Monitoring level	Objective
Smart electricity meters	Critical machine embedded in electricity system	Machine/workshop/factory level	Track the electricity information, including electric quantity, current, voltage, frequency, etc.
Smart water meters	Critical machine embedded in water system	Machine/workshop/factory level	Track the water information, including hydraulic pressure, flow, temperature, etc.
Smart gas meters	Critical machine embedded in gas system	Machine/workshop/factory level	Track the gas information, including flow, pressure, temperature, etc.
Smart quality meters	Critical machine embedded in coal system	Machine/workshop/factory level	Track the coal information, including quality, volume, etc.
Temperature sensors	Embedded in machine	Machine level	Track the temperature data of machine, material, production, etc.
pressure sensors	Embedded in machine	Machine level	Track the pressure data of machine, material, production, etc.
RFID readers/Tags	Critical tool/pallet/operators /AGV/robot arm	Machine level	Track and trace real-time information of manufacturing process, etc.

An overall architecture of energy big data perception and acquisition is designed in Fig. 4. The configurations of various IoT devices in Table 2 are the foundations for collecting the multi-source and heterogeneous energy data.

During the whole production process, IoT devices are deployed for manufacturing resources and energy control points in factory level, workshop level, and machine level, respectively, as shown in Fig. 4. For example, smart meters (electricity meters, gas meters, water meters and quality meters) are used to monitor and capture energy data (electricity, fuel gas, water, and coal) during the production process. Then through the standard communication protocols, such as RS-485/232, Ethernet, Wi-Fi, Bluetooth, Infrared, Modbus, Fieldbus technologies, the captured real-time and non-real-time energy data is transmitted to the enterprise databases.

However, the key measure parameters in the production process face the challenge of collecting energy data in extreme conditions, namely, these parameters cannot be directly measured through IoT devices. Thanks to the soft sensor theory, the soft sensor approach can be used to estimate these key measure parameters by a mathematical relationship between the key measure parameters and the easy-to-measure parameters. The estimation accuracy and reliability of soft sensor have been validated by many research (Li et al., 2015). For example, a simulation result is employed to demonstrate the effectiveness and capability of the soft sensor method, and the comparative studies have demonstrated that the soft sensor approach is much of higher prediction accuracy and reliability than others (Bidar et al., 2017).

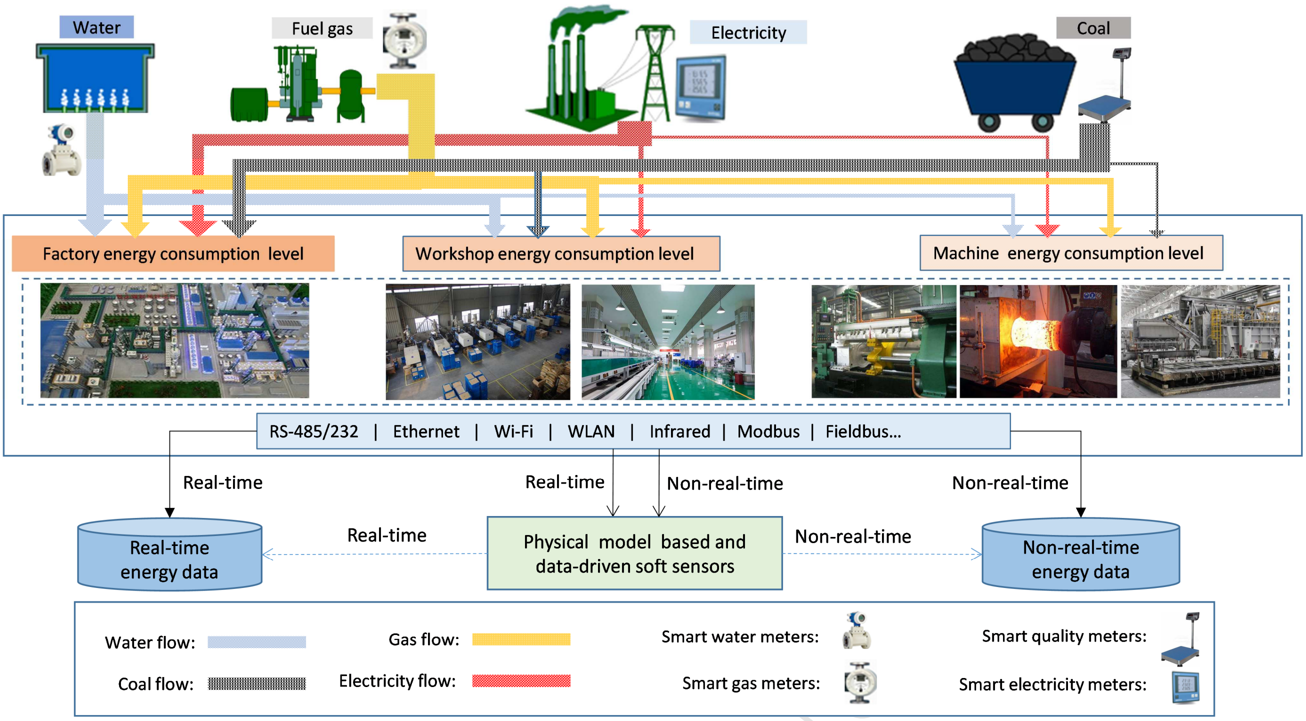


Fig. 4. Overall architecture of energy big data perception and acquisition.

In the designed architecture, the energy data in extreme conditions can be measured indirectly by the other available real-time and non-real-time energy data. The physical model can be created to estimate the key measure variables by the easy-to-measure process variables and the relationship between two types of variables (Kadlec et al., 2009). For example, the real-time power of electrical machinery is equal to the product of real-time electrical current and real-time electrical voltage (Hambley, 2014). So the real-time power can be measured indirectly by the real-time and easy-to-measured electrical current and voltage. Further, the data-driven model can be established through advanced analysis techniques, which provide a powerful modeling toolbox of driven models. Based on the measured energy data, the essential information in extreme conditions will be exploited through the machine learning methods including artificial neural networks, multivariate statistics, fuzzy logic, support vector regression, Gaussian regression, hybrid methods, and so on (Yan et al., 2017).

4.2 Energy big data mining

The closed-loop structure consists of the data layer, model layer, and objective layer, as shown in Fig. 5.

The data layer includes multi-source and heterogeneous energy data, which has three dimensions of energy data, namely the spatial dimension (machine, workshop, factory, system), energy dimension (coal, fuel gas, petroleum, electricity), and time dimension (minute, hour, day, month). According to different application objectives, all of these energy data sets have been cleansed, integrated, and stored in different enterprise databases.

Model layer mainly refers to the models of energy data mining, including energy consumption analysis models (e.g. the input-output process model, the green schedule model, e-p analysis, c-g analysis) and big data mining models (e.g. decision trees, rough set theory, artificial neural networks, apriori algorithms, genetic algorithms, Bayesian estimation) (Wu et al., 2014). An energy data mining model comes from the combination of energy consumption analysis model and data mining model or one of them. On one hand, from energy consumption analysis to energy data mining, for example, the e-p analysis model can be used to find the energy-intensive equipment in the manufacturing process. Then a suitable method of data mining will be selected to eliminate, improve and make use of the energy-intensive equipment reasonably.

On the other hand, from energy data mining to energy consumption analysis, a method of the decision tree can be used to find the bottleneck problem of energy consumption in production and the input-output process model may be selected to solve the problem further. According to different model demands, suitable energy data from three dimensions will be extracted from the data layer.

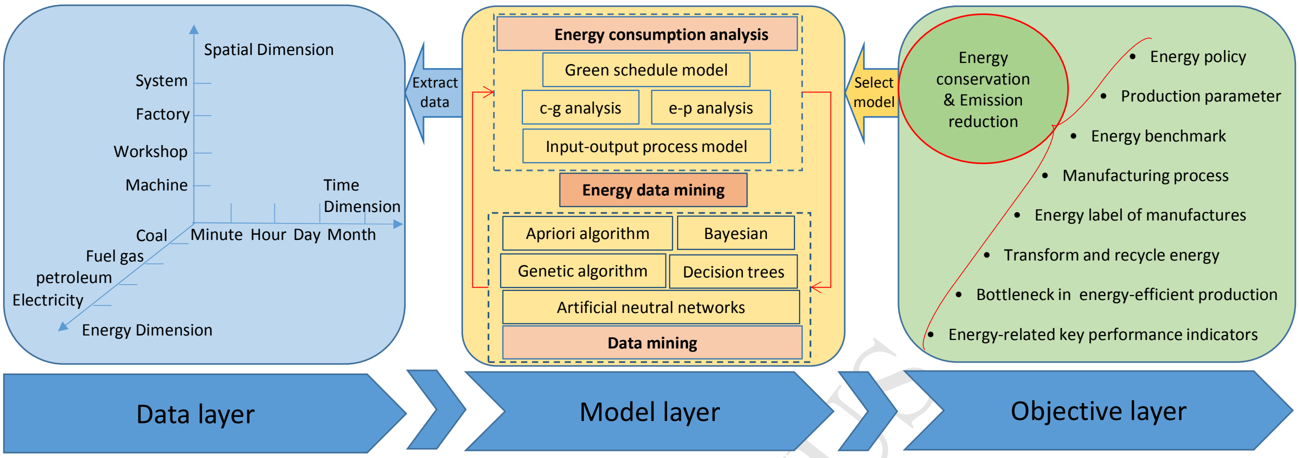


Fig. 5. Closed-loop structure of energy big data mining.

The objective layer is known as an objective set of energy conservation and emission reduction, which comes from the production applications. Based on the literature review of energy management in manufacturing (May et al., 2017), in this research, the production application of the objective layer mainly includes energy policy, production parameter, energy benchmark, manufacturing process, energy label of manufactures, transform and recycle energy, bottleneck in energy-efficient production, energy-related key performance indicators (e-KPIs). These objectives come from all sides principally in macroscopic and microscopic levels. The optimization objective may be one or more of the objective layer. According to different demands of the objective layer, suitable energy data mining model and energy consumption data are selected to carry out the knowledge discovery. For example, on one hand, from the macroscopic level, an energy policy is made by selecting the related model, which may be the input-output process model. On the other hand, from the microscopic level, production parameter can be optimized by selecting the mining algorithm, which may be the genetic algorithm. Finally, based the mined results, the enterprise managers can make better energy-efficient decisions.

The analysis above shows that the closed-loop structure starts from the application objectives, and finally meets the application objectives. Firstly, the application objectives are proposed. Secondly, based on the different objectives, the adaptive models are selected and established. The rule base of selecting model may have the abilities of self-learning, store-memory, and evolution (Sultan and Ahmed, 2017). Thirdly, suitable energy consumption data is extracted to implement energy data mining. Finally, the information and knowledge are obtained to meet the application objectives.

5. A study of application scenario

This section describes a proof-of-concept application scenario to demonstrate how to implement the presented BDDAF for EIMIs. The ceramic manufacturing industries are one of the typical EIMIs with high energy consumption and large emission. The ceramic belongs to MNMP (Li et al., 2014) and the MNMP belongs to six major EIIs (China Economic and Social Development Statics Bulletin, 2010). So the ceramic manufacturing industries are one of EIMIs. The authors have surveyed a partner ceramic manufacturing company (“company X") in Guangdong, China. Company X yearly consumes more than 25 million kWh of electricity, 14 million cubic meters of natural gas, 280 tons of diesel and 350,000 tons of water (Li et al., 2017). Thus, it is suitable to be used to verify the proposed framework.

In company X, based on energy management system, the real-time energy data from raw materials to products can be collected and monitored in the ceramic manufacturing process. The collected multi-source and heterogeneous data are characterized by high volume, high velocity, high variety, and high value, which belong to a typical family of energy big data. Based on energy big data analytics, the managers will find out the unreasonable energy consumption because of abnormal ball mill, production schedule and so on.

5.1 Case description

In the ceramic manufacturing industry, the production chain mainly consists of ball-milling, slip casting, glazing and sintering (Li et al., 2017). Firstly, in the preparation section, raw materials and auxiliary materials are mixed and billed into the slurry with certain moisture and fitness by sifting and de-ironing. Secondly, in the shaping section, the slurry is formed into the semi-finished product by slip casting. Thirdly, the glaze is sprayed on the face of semi-finished product with good decorative effect. Finally, after the sintering, the product is finished, classified, and packaged for storage and selling.

Electricity needs to be used throughout the entire manufacturing process. In order to analyze the most electricity consumption process, then the e-p analysis model (Liu et al., 2006b) of the ceramic product is created as follows.

$$E = \sum_{i=1}^{n} e_i p_i \tag{1}$$

E : The overall electricity consumption per ton of the ceramic product (KJ/ton).

e_i : The electricity intensity of unit process i per ton of the ceramic product (KJ/ton).

p_i : The product ratio of the unit process i, which is the product yield of this unit process.

$(e_i p_i)$: The electricity consumption of the unit process i per ton of the ceramic product (KJ/ton).

$e_1 p_1$: The electricity consumption of the ball mill process per ton of the ceramic product (KJ/ton).

Compared with the energy data of unit process based on the energy management system, it is observed that

$$e_1 p_1 > e_i p_i,\ i = 2,3,...,n \tag{2}$$

That is to say, the ball milling process is the biggest electricity consumption process. All ball mills are in pulp workshops. Therefore we will make a deep analysis of ball mills in a pulp workshop of company X, seeking out the existing problems of energy waste, and propose solutions to save energy and reduce the production cost.

5.2 Energy data acquisition of ball mills

The pulp workshops are divided into old workshop and new workshop. There are 18 8-tons ball mills in the old workshop and 6 40-tons ball mills in the new workshop. The number of ball mills in the old workshop is more than that in the new workshop. Therefore, the total status of ball mills is relatively complex and diverse. Given that, we select the old workshop as the research object. The ball mill is a type of machine used to blend and grind materials of ceramics. Through smart quality meters and smart water meters, materials and water can be mixed in certain weight proportion. The basic parameters (such as the feeding formula and feeding volume) of these ball mills are the same. All ball mills are in use.

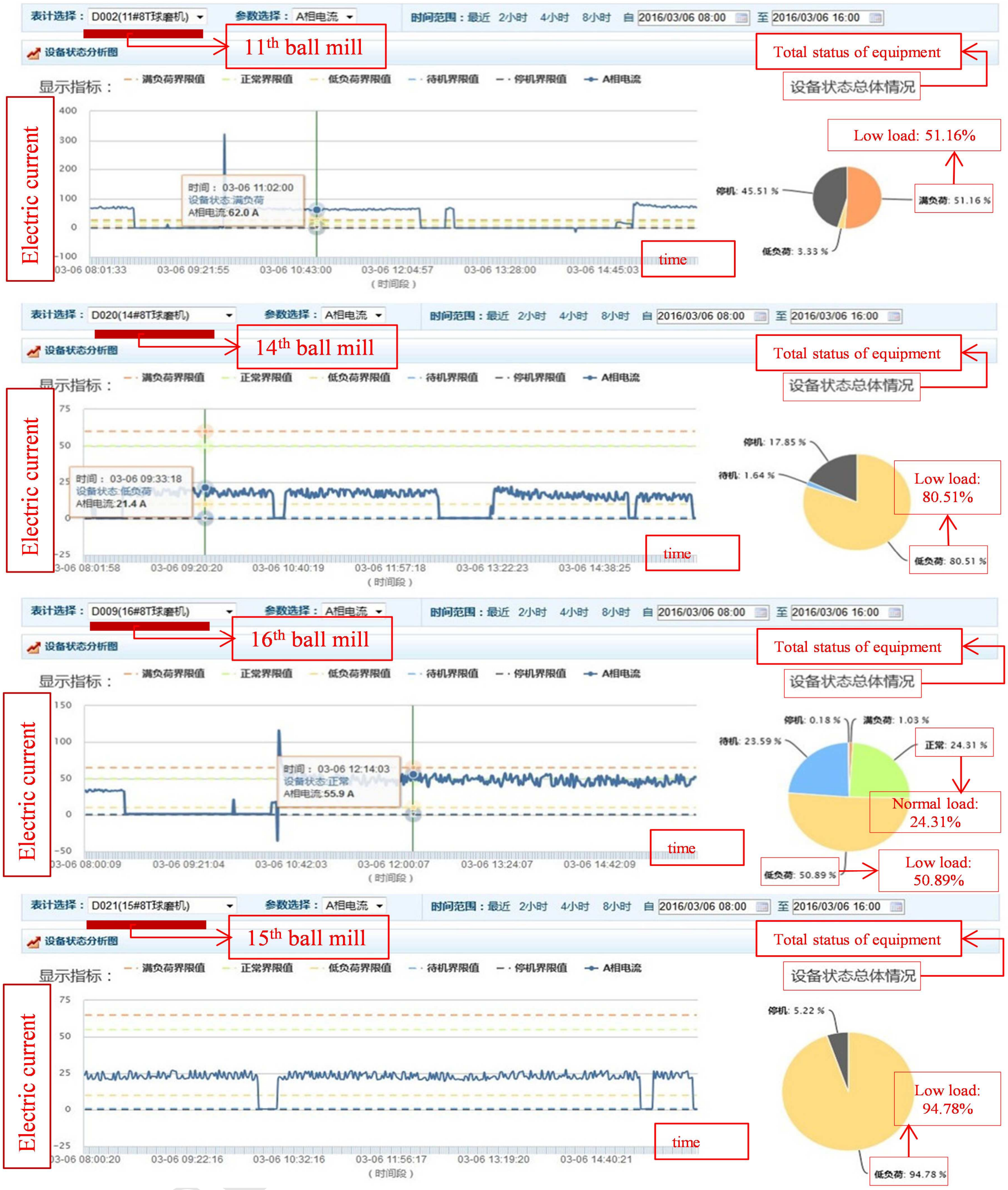


Fig. 6. The status of four ball mills from 08:00 to 16:00 on March 6, 2016 (This is the original screenshot extracted from the case company and thus only available in Chinese)

According to the operation status of 18 8-tons ball mills in the old workshop, they could be roughly divided into four categories, i.e. the unstable (open and close frequently) ball mill, low load ball mill, full load ball mill, and formal load ball mill. The 14th ball mill, 15th ball mill, 16th ball mill, and 11th ball mill belong to the four categories respectively. Other ball mills may overlap between categories. The four more representative ball mills (11th ball mill, 14th ball mill, 15th ball mill and 16th ball mill) are specifically analyzed. The main motor power of the four ball mills is 75kW, 55kW, 75kW and 75kW respectively. Then the main parameters and state of the four balls are analyzed as follows.

Table 3

Main parameters of four ball mills.

Parameters	11th ball mill	14th ball mill	15th ball mill	16th ball mill
Main motor power (kW)	75	55	75	75
Grinding time (hours)	13	17	13	17
Power consumption (kWh)	453	237	225	302
Load status	Full	Low	Low	Normal/Low

Energy consumption of ball mills can be seen from real-time power. However, the real-time power cannot be easily measured. We use the method of soft sensor to computer the relationship between real-time electronic current and real-time power (Hambley, 2014). The real-time power is

$$p(\mathrm{t}) = 3V_{rms}(t)I_{rms}(t)\cos(\theta) \tag{3}$$

Where $p(\mathrm{t})$ is power at time t, $V_{rms}(t)$ is the root-mean-square (rms) line-to-neutral voltage at time t, which is a constant, $I_{rms}(t)$ is the rms line current at time t, and θ is the angle of the load impedances, which is a constant.

Then

$$p(\mathrm{t}) = C_0 I_{rms}(t) \tag{4}$$

Where $C_0 = 3V_{rms}(t)\cos(\theta)$ is a constant.

The real-time power $p(\mathrm{t})$ corresponds to the real-time electronic current $I_{rms}(t)$, so we can study the status of ball mills by the real-time electronic current $I_{rms}(t)$ instead of $p(\mathrm{t})$ in next section.

5.3 Energy data mining of ball mills

The energy management system shows that the ball grinding time of 11th and 15th ball mill is about 13 hours and the ball grinding time of 14th and 16th ball mill is about 17 hours, on March 6, 2016. Meanwhile, the power consumption of 11th, 14th, 15th, and 16th ball mill is 453kWh, 237kWh, 225kWh and 302kWh respectively. From the Fig. 6, we know the 11th ball mill is under full load, the 14th and 15th ball mill are under low load, and 16th ball mill is under normal or low load. The detailed parameters are summarized in Table 3.

5.3.1 The analysis of 15th and 11th ball mill

As shown in Fig. 6, the electric current of 15th ball mill is most stable. Meanwhile, 15th ball mill is under low load and its power consumption (225kWh) is lowest. As a result, the energy efficiency of 15th ball mill is the highest of the four ball mills.

The 11th ball mill is under full load and its power consumption of 11th ball mill is 453kWh, which is around twice than that of 15th ball mill. The accurate analysis of the 11th ball mill and improving its efficiency will achieve some degree of energy conservation. For example, too many or too heavy grinding balls may affect energy efficiency of ball mill. So operators can detect the parameters of ball mill.

5.3.2 The analysis of 14th ball mill

As shown in Fig. 6, the ball grinding state of the 14th ball mill is not stable. It is stopped or standby 4 times. Then, from the energy management system, we find the further status of 14th ball mill from 16:00 March 6 to 0:00 March 8. Furthermore, the stop time of the 14th ball mill is long. The status may mean that the mill needs to be stopped to arrange the materials, which may be related to the production schedule. If it is the case, the managers can arrange the downtime in peak power period, make full use of standard and valley power period to start up, and reduce the electricity cost (Fernandez et al., 2013).

In addition, the motor of the 14th ball mill is an energy-saving motor of 55kW. The power consumption is almost same than that of 15th ball mill, but the ball grinding time is more 4 hours than that of 15th ball mill. According to the field measurements, the 14th ball mill takes about 44.93 seconds every 10 turns, while 15th ball mill takes about 37.98 seconds. This may be caused by the slow speed of the motor. The managers can replace the energy-saving motor of 55kW with an energy-saving motor of 75kW, which can reduce the ball grinding time, stabilize the ball grinding efficiency and reduce the energy consumption.

5.3.3 The analysis of 16th ball mill

When the 16th ball mill is running, the ball mill is under normal or low load, but the ball grinding time is long and the power consumption (302kWh) is large. The possible reason is that the ball mill's grinding speed is slow. The motor load is normal, so the motor problem is eliminated. It is possible that the belts are loose.

Fig. 7. The electric current of 16th ball mill (This is the original screenshot extracted from the case company and thus only available in Chinese)

In addition, the electric current curve of 16th ball mill has a remarkable characteristic. When the materials are put into grinding machine, the ball mill is under normal load. However, after a period of time, the motor is under low load, as shown in Fig. 7. The materials are made of sand and minerals, which is hard to grind. After the feedstock, the electric current is slightly large. Then the solid powder becomes pulp gradually. The friction in pulp reduces and the fluidity is good, resulting in the decrease of the motor load. This situation exists, but if this is the case, then other ball mills should be the same. Nevertheless, this situation exists only in the 16th ball mill.

Fig. **8**. The comparison of belts of 15th and 16th ball mill

We go to the scene to analyze the 16th ball mill. Then we find the situation, as shown in Fig. 8. The grinding speed of 15th and 16th ball mill is measured. They are about 37.98 seconds and 38.68 seconds every 10 turns respectively. It means that the speed of the 15th ball mill is faster than that of 16th ball mill. Furthermore, the 15th ball mill has 18 belts. The belts are in good condition and the tightness is normal. However, there are only 11 belts in the 16th ball mill. The condition and tightness of belts are also OK. The belts of 16th ball mill are few, which may reduce the ball grinding speed. As a result, the ball grinding time increases when the same materials are used to achieve the same physical performance. It is suggested that the increase of the belts of 16th ball mill can reduce the ball grinding time and reduce the power consumption of ball grinding.

In practice, problems of unreasonable or wasteful energy consumption cannot be found easily due to the lack proactive maintenance. As a result, the problem of the 16th ball mill lacking 7 belts is not found by operators, but is analyzed by energy big data mining. The problem of lacking belts may bring the great loss for the company. In detail, as is shown in Table 3, compared to the 15th ball mill, the 16th ball mill wastes the power consumption of 77 kWh and the grinding time of 4 hours because of lacking belts. In other words, timely maintenance and replacement can also improve the energy efficiency of the machine.

5.4 Results

In company X, based on energy management system, the real-time energy consumption of the ball mill in a production cycle can then be estimated based on the real-time electronic current and real-time electronic voltage. The quantitative and qualitative analysis of ball mills is excavated to find that 11th ball mill has places to save energy, 14th ball mill can reduce the electricity cost by production schedule, 16th ball mill needs to be repaired by adding belts. According to energy big data analytics, the ball grinding efficiency of 15th ball mill is highest of four ball mills. The various parameters of the 15th ball mill should be analyzed in detail, which can be extended to other balls and effectively reduce energy consumption and costs. In addition, the rules and regulations are also made to improve maintenance of the ball mills and the daily maintenance. For example, according to the energy consumption ranking of the ball mill of unit product, the managers can optimize the production schedule and maintain low-energy-consumption equipment timely.

The comprehensive energy consumption of company X is 21,000 tons of SCE in 2014. After the implementation of the proposed framework, based on the energy management system, it has been possible to monitor and analyze the multi-source and heterogeneous energy data generated during the whole production process. Under the proposed framework, problems that were previously overlooked have been discovered, and it has been confirmed by energy big data analytics that the failure of even just a few broken belts will have a serious impact on energy consumption. Then energy consumption of unit product has been reduced by 3% one month later and energy costs are saved about 4% after half a year in June 2015 (Foshan Dingxing Technology Company, 2016a).

6. Discussions

6.1 Effectiveness of the proposed framework in theory

Energy conservation and emission reduction have grown up to be important and effective means for energy-intensive manufacturers to improve their competitiveness. Facing various modes of energy conservation and emission reduction in practice, manufacturers have to choose their optimal modes (Ouyang and Shen, 2017). Manufacturing industries have been able to reduce energy consumption and waste in their production process by adopting advanced production management paradigms (e.g. lean and Six Sigma programs) (Auschitzky et al., 2014). However, in certain production environments, especially in EIMIs, high energy consumption and high pollution are a fact of life, sometimes even after advanced production management paradigms have been applied. Therefore, EIMIs need a new systematic and integrated approach to save energy and reduce emission. Energy big data analytics in manufacturing provides just such an approach, which is a new interdisciplinary research area of energy, big data, and manufacturing. With the background of energy big data in manufacturing, a BDDAF for EIMIs is proposed. The proposed architecture framework provides approaches, techniques, tools, rules, principles, and practices for data acquisition, storage, preprocessing, mining, decision-making, and application. In more details, the effectiveness of the proposed framework is as follows.

- A major gap in literature addressed by this paper has been the lack of a systematic and comprehensive architecture for EIMIs, the proposed framework has been found effective to fill in the gap. Energy big data analytics refers to the application of statistics and other mathematical tools to analyze manufacturing energy consumption for optimizing production process, improving material and energy efficiency, reducing emissions, and saving costs. Therefore, EIMIs taking advantage of energy big data analytics can implement CP strategy. Research and development of energy big data analytics and applications have brought new opportunities for understanding manufacturing energy consumption behavior, which could provide a theoretical and practical research direction in the academic and industrial field.
- In EIMIs, a deep insight into the energy consumption patterns during the whole process is a prerequisite for improving the material and energy efficiency. However, the measurement of energy consumption faces challenges due to the complexity that arises from the variety of energy use in harsh production environment. Therefore, the overall architecture of energy big data acquisition is designed to sense the multi-source and heterogeneous energy big data during the various processes, each one having their unique energy consumption characteristics. In the designed architecture, the energy data in extreme conditions can be estimated by the physical model based and data driven soft sensors.
- The production chain consists of continuous flow and discrete flow manufacturing processes in EIMIs. The existing approaches cannot be used to model the energy consumption patterns due to the more complex energy consumption characteristics and process interactions. Based on the collected energy big data in manufacturing, the closed-loop structure of energy big data mining, which is integrated into the methods of the energy consumption analysis, is designed to improve energy efficiency and promote energy conservation. The problem of energy waste can be found

by mining energy big data. For example, in the case study, according to the object of optimizing the material and energy efficiency in the manufacturing process, the e-p analysis model is selected and then electricity consumption data in every machining process are extracted to verify that the ball milling process is the biggest electricity consumption process. Further, the quantitative and qualitative analysis of ball mills is mined to find out the obvious and potential problem of energy waste.

6.2 Effectiveness of the proposed framework in practice

Section 5 describes a proof-of-concept application scenario to demonstrate the proposed framework for EIMIs. In practice, all of our partner companies have adopted the proposed framework based on energy management system and have saved the energy and cost. For example, Company A is one of the aluminum top ten enterprises in China. Through the monitoring function of energy network, energy leakage can be detected in time to reduce energy waste. Based on energy efficiency analysis, we can find out the unreasonable operation of air compressor and boiler, prevent the recurrence of the similar situation, and improve energy efficiency. In the second half of 2015, the energy saving rate reached 4% (Foshan Dingxing Technology Company, 2016b).

Company C is an energy-intensive enterprise of manufacturing copper tube. The company mainly consumes electricity. In 2012, electricity cost is more than CNY 100 million. It is used to optimize off-peak power consumption of energy-intensive processes, such as melting casting, annealing. In the fourth quarter of 2014, the saving cost of energy was about CNY 500 thousand (Foshan Dingxing Technology Company, 2016c).

Table 4

Application effect of our partner manufacturing companies

Manufacturing company	Types of company	Annual energy consumption	Online time	Application effect	Data source
Company A	Aluminum profile	13 thousand tons of SCE in 2014	July in 2015	Saving the energy about 4 % in the second half of 2015	(Foshan Dingxing Technology Company, 2016b)
Company C	Copper tube	More than CNY 100 million of electricity cost in 2012	August in 2014	Saving the energy cost about CNY 500 thousand in the fourth quarter of 2014	(Foshan Dingxing Technology Company, 2016c)
Company R	Rubber tire	75 million kWh of electricity and 55000 tons of coal in 2012	July in 2012	Saving about CNY 3.4 million of electricity cost in 2013	(Foshan Dingxing Technology Company, 2016d)

Company R is one of the largest tire production bases in south China, which is a typical energy-intensive manufacturing enterprise. The system collects and analyzes all kinds of data related to energy, finds out the abnormal energy consumption in time, and helps the management personnel to make the production adjustment in time. The energy efficiency assessment of the equipment and team is realized, and the energy and other resources can be saved to the maximum extent under the premise of ensuring the normal production. In 2013, the saving electricity cost is about CNY 3.4 million (Foshan Dingxing Technology Company, 2016d).

The application effect of our partner manufacturing companies above is compared, as shown in Table 4.

6.3 Limitations

The proposed framework and key enabling technologies for energy big data analytics provide a new kind of infrastructure to improve energy efficiency in the whole production process. However, the soft sensor technologies are only proposed in architecture, lacking implementation in detail. In addition, energy data mining models should be studied deeply integrating energy consumption analysis and data mining. Future research can be carried out in the following aspects. Firstly, how to use the soft sensor approaches to capture real-time energy data in harsh production environment. Secondly, by using the data mining theory, a mathematical model will be established to identify the hidden knowledge and rules from the multi-source and heterogeneous energy big data. For example, how to establish a quantitative model such as state space equations and real-time distributed control for energy efficient manufacturing systems.

7. Managerial implications

Managerial implications could be generated from hidden patterns, associated relationships and key findings of energy big data, which are useful when various department managers are making energy-efficient decisions accordingly. The rest of the section describes implications to assist managers in EIMIs to make the energy-efficient decisions from the government department, the production department, and the research & development department.

7.1 The government department

Firstly, the government can achieve the energy-saving targets of EIMIs by using big data analytics about the industrial structure, product structure, enterprise structure, and production scale. When the industrial structure is developed to the direction of energy-saving type, EIMIs can use big data analytics to analyze the data related to energy consumption during the whole product lifecycle stage such as design, production, distribution, usage, maintenance, reuse, and remanufacturing (Y. Zhang et al., 2017b). Then EIMIs can try to optimize the factors that have the greatest effects on energy consumption in the production stage. As a result, the products will be sustainable and the energy consumption of products will be reduced. In addition, the structure and organization of production, the process and technology, and the mode of energy usage have a great influence on energy consumption in EIMIs, especially in small and medium-sized EIMIs. Therefore, the government should organize these small and medium-sized EIMIs together to realize the goal of efficient energy conservation by joint management.

Secondly, the government can design relevant energy policies and standards with regard to the benchmark rating system through big data analytics of similar EIMIs. When the energy consumption of the EIMIs is below the grade of the energy benchmarking system (Cai et al., 2017), the firm should be subjected to the financial and administrative penalty in some degree. Incentive schemes can also be implemented for EIMIs that satisfy the grade of the energy benchmarking system.

7.2 The production department

Firstly, the machine spends large amounts of time in the standby and idle states because of the poor energy consciousness of operators, resulting in a massive waste of energy. The process parameters mainly depend on the subjective consciousness and the machining experience of operators to satisfy the machining requirements but ignore energy-consumption issues.

Based on big data analytics, operators can monitor the real-time energy consumption, search for reasons for the exceptions, adjust the process parameters and realize green scheduling (May et al., 2017) including scheduling models and algorithms for energy efficiency, multi-objective optimization (e.g. makespan and energy) and so on. In addition, it is necessary to remind timely the maintenance of the equipment through the historical data mining to reduce unnecessary waste of energy.

The e-KPIs for improving energy efficiency can be used to strengthen the theoretical base necessary to support energy-based decision-making in EIMIs (May et al., 2015). With the help of big data analytics, the method of improving e-KPIs can identify firm-specific energy drivers in their production system and make the energy behavior profile of the production system transparent. For example, the workshop managers can find the bottleneck problem of energy consumption in production and solve it through big data analytics.

7.3 The research & develop department

The input of research and development of EIMIs should be increased in the future. Inputs in research and development may stimulate the emergence of new technologies which can produce high standards and environmentally friendly products of EIMIs (Lin and Tan, 2017). Based on big data analytics, the research & development department can save energy indirectly by renewing the equipment, improving the level of technological processes and the level of operators, and transforming as well as recycling energy.

As time goes by, more and more energy data for mass production in machining system are generated, big data analytics can be used to analyze and mine these data. The transformation and usage of energy are working in equipment level, so the quality and performance of equipment are determined to a large extent to effective usage of energy. It is very important to eliminate, improve and make use of the energy-intensive equipment reasonably. Based on big data analytics, the optimization of the process parameter and operator parameter can be obtained. In addition, the waste of energy, such as combustible surplus energy and thermal residual energy, can be recycled by cleaner technologies.

8. Conclusions

Currently, for traditional EIMIs, it is difficult to collect the multi-source and heterogeneous energy big data in harsh production environment. In addition, it is difficult to discover and excavate hidden knowledge from energy big data to avoid unnecessary wastes, and then to provide valuable energy-efficient knowledge for managers during implementation of the CP strategy.

To address these challenges, this article presented a BDDAF for EIMIs. Several contributions were important in this research. The first contribution was the BDDAF and its key components. Under the framework, manufacturing enterprises can obtain a new mode and analytical approach during the whole process for reducing waste, emissions, and costs. The second contribution was the architecture of energy big data perception and acquisition. Using the architecture, the multi-source and heterogeneous energy big data are collected by using the IoT technologies and soft sensor approaches in harsh production environment. The third contribution was the closed-loop structure of energy big data mining, which was put forward to mine valuable knowledge and patterns from the multi-source and heterogeneous energy data in EIMIs. Furthermore, the fourth contribution was the summarized managerial implications from the perspective of the government department, the production department, and the research & development department. Then three departments are able to make energy-efficient decisions in different solutions, as well as effectively promote the implementation of CP strategy.

Moreover, the ultimate goals of BDDAF are energy conservation and emission reduction, which may benefit the manufacturing enterprises, the government, and the society. Firstly, improving energy efficiency in manufacturing can be considered as a pragmatic and an attractive solution, because it assists manufactures to reduce their production cost, ultimately enhancing their sustainable competitive advantage in the market. Secondly, the manufacturing industries obtain continuously targets of the energy conservation and emission reduction, thereby helping the government to achieve the target of the Paris Agreement (United Nations Framework Convention on Climate Change, 2015). Finally, the threat to human health (e.g. lung diseases, respiratory diseases, immune diseases) will be reduced continually due to reduction of the pollutants (e.g. sulfur oxides and nitrogen oxides).

Acknowledgement

The authors would like to thank the Faenza Ceramics Group for providing the data sets for this work and given some useful information on their utilization. The authors also thank Dr. Paolo Massioni for his suggestions and the fruitful discussions. The authors gratefully acknowledge financial supports of National Science Foundation of China (51675441, 51475096, U1501248), and the Fundamental Research Funds for the Central Universities (3102017jc04001). This study is also supported in part by the Circularis (Circular Economy through Innovating Design) project funded by Vinnova - Sweden's Innovation Agency (2016-03267) and the Simon (New Application of AI for Services in Maintenance towards a Circular Economy) project funded by Vinnova - Sweden's Innovation Agency (2017-01649).

References

Auschitzky, E., Hammer, M., Rajagopaul, A., 2014. How big data can improve manufacturing. McKinsey Co. Inc. 1–4.

Bidar, B., Sadeghi, J., Shahraki, F., Khalilipour, M.M., 2017. Data-driven soft sensor approach for online quality prediction using state dependent parameter models. Chemom. Intell. Lab. Syst. 162, 130–141. doi:10.1016/j.chemolab.2017.01.004

BP Statistical Review, 2017. China' energy market in 2016. http://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review-2017/bp-statistical-review-of-world-energy-2017-china-insights.pdf (accessed 15 August 2017, in Chinese).

Bunse, K., Vodicka, M., 2010. Managing energy efficiency in manufacturing processes – implementing energy performance in production information technology systems, in: What Kind of Information Society? Governance, Virtuality, Surveillance, Sustainability, Resilience. pp. 260–268. doi:10.1007/978-3-642-15479-9_25

Cai, W., Liu, F., Zhang, H., Liu, P., Tuo, J., 2017. Development of dynamic energy benchmark for mass production in machining systems for energy management and energy-efficiency improvement. Appl. Energy 202, 715–725. doi:10.1016/j.apenergy.2017.05.180

Cai J., 2009. Energy consumption analysis and energy-saving countermeasure research for integrated steelworks. Angang Techniques. 2, 1-6 (in Chinese).

Cattell, R., 2011. Scalable SQL and NoSQL data stores. Acm Sigmod Rec. 39, 12–27. doi:10.1145/1978915.1978919

Chan, D.Y., Huang, C., Lin, W., Hong, G., 2014. Energy efficiency benchmarking of energy-intensive industries in Taiwan. Energy Convers. Manag. 77, 216–220. doi:10.1016/j.enconman.2013.09.027

China Economic and Social Development Statics Bulletin, 2010. China economic and social development statics bulletin. http://www.ce.cn/macro/more/201102/28/t20110228_22253696.shtml (accessed 21 March 2017, in Chinese).

Chong, D., Shi, H., 2015. Big data analytics: a literature review. J. Manag. Anal. 2, 175–201. doi:10.1080/23270012.2015.1082449

Diamantoulakis, P.D., Kapinas, V.M., Karagiannidis, G.K., 2015. Big data analytics for dynamic energy management in smart grids. Big Data Res. 2, 94–101. doi:10.1016/j.bdr.2015.03.003

Fan, L.W., Pan, S.J., Liu, G.Q., Zhou, P., 2017. Does energy efficiency affect financial performance? Evidence from Chinese energy-intensive firms. J. Clean. Prod. 151, 53–59. doi:10.1016/j.jclepro.2017.03.044

Fernandez, M., Li, L., Sun, Z., 2013. "Just-for-Peak" buffer inventory for peak electricity demand reduction of manufacturing systems. Int. J. Prod. Econ. 146, 178–184. doi:10.1016/j.ijpe.2013.06.020

Fong, J., Wong, H.K., Huang, S.M., 2003. Continuous and incremental data mining association rules using frame metadata model. Knowledge-Based Syst. 16, 91–100. doi:10.1016/S0950-7051(02)00076-X

Foshan Dingxing Technology Company, 2016a. Foshan faenza sanitary ware limited companuy. http://www.dingx.cn/successstories-fe.html (accessed 1 April 2017, in Chinese).

Foshan Dingxing Technology Company, 2016b. Guangdong jianmei aluminum profile limited companuy.

http://www.dingx.cn/successstories-jm.html (accessed 26 April 2018, in Chinese).

Foshan Dingxing Technology Company, 2016c. Guangdong zhuhai longfeng precision copper tube limited companuy. http://www.dingx.cn/successstories-lf.html (accessed 26 April 2018, in Chinese).

Foshan Dingxing Technology Company, 2016d. Guangzhou fengli rubber tire limited companuy. http://www.dingx.cn/successstories-fllt.html (accessed 26 April 2018, in Chinese).

Fournier-Viger, P., Chun, J., Lin, -Wei, Kiran, R.U., Koh, Y.S., Thomas, R., 2017. A survey of sequential pattern mining. Ubiquitous Int. 1, 54–77.

Haller, S., Karnouskos, S., Schroth, C., 2009. The Internet of things in an enterprise context, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 14–28. doi:10.1007/978-3-642-00985-3_2

Hambley, A.R., 2014. Steady-state sinusoidal analysis, in: Electrical Engineering: Principles & Applications (Fourth Edition). China Machine Press, Beijing, pp. 206-207.

Huang, Y., Luo, J., Xia, B., 2013. Application of cleaner production as an important sustainable strategy in the ceramic tile plant-a case study in Guangzhou, China. J. Clean. Prod. 43, 113–121. doi:10.1016/j.jclepro.2012.12.013

Jacobs, A., 2009. The pathologies of big data. Commun. ACM 52, 36–44. doi:10.1145/1536616.1536632

Jian, W., Zhu, L., Xu, Z., Chen, X., 2017. A variable selection method for soft sensor development through mixed integer quadratic programming. Chemom. Intell. Lab. Syst. 167, 85–95. doi:10.1016/j.chemolab.2017.05.011

Kadlec, P., Gabrys, B., Strandt, S., 2009. Data-driven soft sensors in the process industry. Comput. Chem. Eng. doi:10.1016/j.compchemeng.2008.12.012

Kara, S., Bogdanski, G., Li, W., 2011. Electricity metering and monitoring in manufacturing systems, in: Glocalized Solutions for Sustainability in Manufacturing - Proceedings of the 18th CIRP International Conference on Life Cycle Engineering. Springer, Berlin, pp. 1–10. doi:10.1007/978-3-642-19692-8-1

Kjaerheim, G., 2005. Cleaner production and sustainability. J. Clean. Prod. 13, 329–339. doi:10.1016/S0959-6526(03)00119-7

Koseleva, N., Ropaite, G., 2017. Big data in building energy efficiency: understanding of big data and main challenges. Procedia Eng. 172, 544–549. doi:10.1016/j.proeng.2017.02.064

Lechtenböhmer, S., Nilsson, L.J., Åhman, M., Schneider, C., 2016. Decarbonising the energy intensive basic materials industry through electrification-implications for future EU electricity demand. Energy 115, 1623–1631. doi:10.1016/j.energy.2016.07.110

Lee, W., On, B.-W., Lee, I., Choi, J., 2014. A big data management system for energy consumption prediction models. Ninth Int. Conf. Digit. Inf. Manag. (ICDIM 2014) 156–161. doi:10.1109/ICDIM.2014.6991404

Li, H., Yang, H., Yang, B., Zhu, C., Yin, S., 2017. Modelling and simulation of energy consumption of ceramic production chains with mixed flows using hybrid Petri nets. Int. J. Prod. Res. 1–18. doi:10.1080/00207543.2017.1391415

Li, L., Wang, J., Tan, Z., Ge, X., Zhang, J., Yun, X., 2014. Policies for eliminating low-efficiency production capacities and improving energy efficiency of energy-intensive industries in China. Renew. Sustain. Energy Rev. 39, 312–326. doi:10.1016/j.rser.2014.07.099

Li, W., Wang, D., Zhou, X., Chai, T., 2015. An improved multi-source based soft sensor for measuring cement free lime content. Inf. Sci. 323, 94–105. doi:10.1016/j.ins.2015.06.035

Lin, B., Tan, R., 2017. Estimating energy conservation potential in China's energy intensive industries with rebound effect. J. Clean. Prod. 156, 899–910. doi:10.1016/j.jclepro.2017.04.100

Lin, B., Tan, R., 2016. Ecological total-factor energy efficiency of China's energy intensive industries. Ecol. Indic. 70, 480–497. doi:10.1016/j.ecolind.2016.06.026

Lin, C.W., Hong, T.P., Lu, W.H., 2009. The Pre-FUFP algorithm for incremental mining. Expert Syst. Appl. 36, 9498–

9505. doi:10.1016/j.eswa.2008.03.014

Lin, J.C.W., Yang, L., Fournier-Viger, P., Hong, T.P., Voznak, M., 2017. A binary PSO approach to mine high-utility itemsets. Soft Comput. 21, 5103–5121. doi:10.1007/s00500-016-2106-1

Liu, L., Aye, L., Lu, Z., Zhang, P., 2006a. Effect of material flows on energy intensity in process industries. Energy 31, 1870–1882. doi:10.1016/j.energy.2005.07.003

Liu, L., Aye, L., Lu, Z., Zhang, P., 2006b. Analysis of the overall energy intensity of alumina refinery process using unit process energy intensity and product ratio method. Energy 31, 1167–1176. doi:10.1016/j.energy.2005.04.013

Liu, W., Wang, Z., 2017. The effects of climate policy on corporate technological upgrading in energy intensive industries: Evidence from China. J. Clean. Prod. 142, 3748–3758. doi:10.1016/j.jclepro.2016.10.090

Liu, Y., 2013. Sustainable competitive advantage in turbulent business environments. Int. J. Prod. Res. 51, 2821–2841. doi:10.1080/00207543.2012.720392

Liu, Y., Liang, L., 2015. Evaluating and developing resource-based operations strategy for competitive advantage: an exploratory study of Finnish high-tech manufacturing industries. Int. J. Prod. Res. 53, 1019–1037. doi:10.1080/00207543.2014.932936

Lv, J., Tang, R., Jia, S., Liu, Y., 2016. Experimental study on energy consumption of computer numerical control machine tools. J. Clean. Prod. 112, 3864–3874. doi:10.1016/j.jclepro.2015.07.040

Lv, J., Tang, R., Tang, W., Jia, S., Liu, Y., Cao, Y., 2018. An investigation into methods for predicting material removal energy consumption in turning. J. Clean. Prod. in press. doi:10.1016/j.jclepro.2018.05.035

May, G., Barletta, I., Stahl, B., Taisch, M., 2015. Energy management in production: A novel method to develop key performance indicators for improving energy efficiency. Appl. Energy 149, 46–61. doi:10.1016/j.apenergy.2015.03.065

May, G., Stahl, B., Taisch, M., Kiritsis, D., 2017. Energy management in manufacturing: From literature review to a conceptual framework. J. Clean. Prod. 167, 1464–1489. doi:10.1016/j.jclepro.2016.10.191

Napp, T.A., Gambhir, A., Hills, T.P., Florin, N., Fennell, P.S., 2014. A review of the technologies, economics and policy instruments for decarbonising energy-intensive manufacturing industries. Renew. Sustain. Energy Rev. 30, 616–640. doi:10.1016/j.rser.2013.10.036

National Bureau of Statistics of China, 2016. China energy statistical yearbook. http://data.stats.gov.cn/english/easyquery.htm?cn=C01 (accessed 8 August 2017).

O'Driscoll, E., O'Donnell, G.E., 2013. Industrial power and energy metering – a state-of-the-art review. J. Clean. Prod. 41, 53–64. doi:10.1016/j.jclepro.2012.09.046

Ouyang, J., Shen, H., 2017. The choice of energy saving modes for an energy-intensive manufacturer considering non-energy benefits. J. Clean. Prod. 141, 83–98. doi:10.1016/j.jclepro.2016.08.142

Polenske, K.R., McMichael, F.C., 2002. A Chinese cokemaking process-flow model for energy and environmental analyses. Energy Policy 30, 865–883. doi:10.1016/S0301-4215(01)00147-1

Porzio, G.F.., Fornai, B.., Amato, A.., Matarese, N.., Vannucci, M.., Chiappelli, L.., Colla, V.., 2013. Reducing the energy consumption and CO2 emissions of energy intensive industries through decision support systems - An example of application to the steel industry. Appl. Energy 112, 818–833. doi:10.1016/j.apenergy.2013.05.005

Rabl, T., Gómez-Villamor, S., Sadoghi, M., Muntés-Mulero, V., Jacobsen, H.-A., Mankovskii, S., 2012. Solving big data challenges for enterprise application performance management. Proc. VLDB Endow. 5, 1724–1735. doi:10.14778/2367502.2367512

Shang, C., Yang, F., Huang, D., Lyu, W., 2014. Data-driven soft sensor development based on deep learning technique. J. Process Control 24, 223–233. doi:10.1016/j.jprocont.2014.01.012

Shrouf, F., Miragliotta, G., 2015. Energy management based on Internet of Things: practices and framework for adoption in production management. J. Clean. Prod. 100, 235–246. doi:10.1016/j.jclepro.2015.03.055

Shvachko, K., Kuang, H., Radia, S., Chansler, R., 2010. The Hadoop distributed file system, in: 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies, MSST2010. pp. 1–10. doi:10.1109/MSST.2010.5496972

Song, C.U., Oh, W., 2015. Determinants of innovation in energy intensive industry and implications for energy policy. Energy Policy 81, 122–130. doi:10.1016/j.enpol.2015.02.022

Song, M., Cen, L., Zheng, Z., Fisher, R., Liang, X., Wang, Y., Huisingh, D., 2017. How would big data support societal development and environmental sustainability? Insights and practices. J. Clean. Prod. 142, 489–500. doi:10.1016/j.jclepro.2016.10.091

Sultan, M., Ahmed, K.N., 2017. SLASH: Self-learning and adaptive smart home framework by integrating IoT with big data analytics, in: 2017 Computing Conference. pp. 530–538. doi:10.1109/SAI.2017.8252147

Tao, F., Wang, Y., Zuo, Y., Yang, H., Zhang, M., 2016. Internet of Things in product life-cycle energy management. J. Ind. Inf. Integr. 1, 26–39. doi:10.1016/j.jii.2016.03.001

Tao, F., Zuo, Y., Xu, L. Da, Lv, L., Zhang, L., 2014. Internet of Things and BOM-based life cycle assessment of energy-saving and emission-reduction of products. IEEE Trans. Ind. Informatics 10, 1252–1261. doi:10.1109/TII.2014.2306771

Thollander, P., Ottosson, M., 2010. Energy management practices in Swedish energy-intensive industries. J. Clean. Prod. 18, 1125–1133. doi:10.1016/j.jclepro.2010.04.011

United Nations Framework Convention on Climate Change, 2015. Paris agreement. http//unfccc.int/files/home/application/pdf/paris_agreement.pdf (accessed 29 August 2017).

Wang, W., Yang, H., Zhang, Y., Xu, J., 2017. IoT-enabled real-time energy efficiency optimisation method for energy-intensive manufacturing enterprises. Int. J. Comput. Integr. Manuf. 0, 1–18. doi:10.1080/0951192X.2017.1337929

Wang, Z., Dazianoc, R.A., Songd, M., Li, S., Zhang, B., Wang, Y., 2017. Call for papers for Special Volume of the Journal of Cleaner Production: Sustainable consumption and big data. https//www.journals.elsevier.com/journal-of-cleaner-production/call-for-papers/call-for-papers-for-special-volume-of-the-journal-of-scbd (accessed 15 April 2017).

Wang, Z., Wang, C., Yin, J., 2015. Strategies for addressing climate change on the industrial level: affecting factors to CO2 emissions of energy-intensive industries in China. Nat. Hazards 75, 303–317. doi:10.1007/s11069-014-1115-6

Wu, X., Zhu, X., Wu, G.-Q., Ding, W., 2014. Data mining with big data. IEEE Trans. Knowl. Data Eng. 26, 97–107. doi:10.1109/TKDE.2013.109

Yan, W., Shao, H., Wang, X., 2004. Soft sensing modeling based on support vector machine and Bayesian model selection. Comput. Chem. Eng. 28, 1489–1498. doi:10.1016/j.compchemeng.2003.11.004

Yan, W., Tang, D., Lin, Y., 2017. A data-driven soft sensor modeling method based on deep learning and its application. IEEE Trans. Ind. Electron. 64, 4237–4245. doi:10.1109/TIE.2016.2622668

Yang, W., Liu, X., Zhang, L., Yang, L.T., 2013. Big data real-time processing based on storm, in: 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications. pp. 1784–1787. doi:10.1109/TrustCom.2013.247

Zaki, M.J., 2000. Parallel and distributed data mining: An Introduction. Genome 1–23. doi:10.1007/3-540-46502-2_1

Zhang, G., Zhang, Y., Xu, X., Zhong, R.Y., 2017. An augmented Lagrangian coordination method for optimal allocation of cloud manufacturing services. J. Manuf. Syst. in press. doi:10.1016/j.jmsy.2017.11.008

Zhang, Y., Guo, Z., Lv, J., Liu, Y., 2018a. A framework for smart production-logistics systems based on CPS and industrial IoT. IEEE Trans. Ind. Informatics in press. doi:10.1109/TII.2018.2845683

Zhang, Y., Qian, C., Lv, J., Liu, Y., 2017a. Agent and cyber-physical system based self-organizing and self-adaptive intelligent shopfloor. IEEE Trans. Ind. Informatics 13, 737–747. doi:10.1109/TII.2016.2618892

Zhang, Y., Ren, S., Liu, Y., Sakao, T., Huisingh, D., 2017b. A framework for Big Data driven product lifecycle management. J. Clean. Prod. 159, 229–240. doi:10.1016/j.jclepro.2017.04.172

Zhang, Y., Ren, S., Liu, Y., Si, S., 2017c. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. J. Clean. Prod. 142, 626–641. doi:10.1016/j.jclepro.2016.07.123

Zhang, Y., Wang, J., Liu, Y., 2017d. Game theory based real-time multi-objective flexible job shop scheduling considering environmental impact. J. Clean. Prod. 167, 665–679. doi:10.1016/j.jclepro.2017.08.068

Zhang, Y., Zhang, G., Wang, J., Sun, S., Si, S., Yang, T., 2015. Real-time information capturing and integration framework of the internet of manufacturing things. Int. J. Comput. Integr. Manuf. 28, 811–822. doi:10.1080/0951192X.2014.900874

Zhang, Y., Zhu, Z., Lv, J., 2018b. CPS-based smart control model for shopfloor material handling. IEEE Trans. Ind. Informatics 14, 1766–1755. doi:10.1109/TII.2017.2759319

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. Int. J. Prod. Econ. 165, 260–272. doi:10.1016/j.ijpe.2015.02.014

Zhong, R.Y., Xu, C., Chen, C., Huang, G.Q., 2017. Big data analytics for physical internet-based intelligent manufacturing shop floors. Int. J. Prod. Res. 55, 2610–2621. doi:10.1080/00207543.2015.1086037

Zhou, K., Fu, C., Yang, S., 2016. Big data driven smart energy management: From big data to big insights. Renew. Sustain. Energy Rev. 56, 215–225. doi:10.1016/j.rser.2015.11.050

Zhou, K., Yang, S., 2016. Understanding household energy consumption behavior: The contribution of energy big data analytics. Renew. Sustain. Energy Rev. 56, 810–819. doi:10.1016/j.rser.2015.12.001