INTERNATIONAL JOURNAL OF LOGISTICS: RESEARCH AND APPLICATIONS
https://doi.org/10.1080/13675567.2020.1870674

Sustainability analysis of supply chain via particulate matter emissions prediction in China

Yelin Qian, Zhenmin Li and Rui Tan

School of Management, Hefei University of Technology, Hefei, People's Republic of China

ABSTRACT

With the increasing importance of air pollutant emissions to the platform economy and green supply chain management, it is essential to analyse the trend and correlation between particulate matter emissions and supply chain statistics. Typical approaches do not integrate particulate matter prediction with the sustainability analysis, and suffer from common issues such as low classification accuracy and unstable prediction performance. In this study, we propose an integrated analytical framework for sustainability analysis of supply chain management through particulate matter emissions prediction. Specifically, we performance trend and correlation analysis between particulate matter emissions ($PM_{2.5}$ and PM_{10}) and supply chain statistics in Beijing of China. We combine the boosting algorithm and neural network method to predict particulate matter emissions. Experimental results show that our prediction model achieved high performance. Sustainability analysis shows that the steady growth of the supply chain operations is accompanied by decreasing air pollutant emissions in China.

ARTICLE HISTORY

Received 17 August 2020
Accepted 21 December 2020

KEYWORDS

Sustainability analysis; supply chain management; particulate matter emission prediction; boosting neural network

1. Introduction

Today, human society is facing very severe environmental and resource problems. The green and sustainable supply chain are the concrete manifestations of sustainable development strategies that comprehensively consider these two issues in platform economy management. In recent years, continued urbanisation and industrialisation, particularly in developing countries, has led to severe deterioration in air quality and a rapid increase in the degree of contamination (Shi and Yu 2020). Air pollution has substantially affected the living environment of the human population and endangered health (Stieb et al. 2019). The United Nations (UN) issued Sustainable Development Goals (SDGs) in 2015, serving as a blueprint for building a more sustainable future across the globe. There are totally 17 goals in SDGs, addressing issues in economic, environmental, and social areas. These goals are interrelated and expected to be fulfilled by 2030 (2015). $PM_{2.5}$ emissions have an adverse effect on public health (Yin, Pizzol, and Xu 2017). At present, many countries attach great importance to inhalable total suspended particulates that cannot be blocked by the human upper respiratory tract, especially inhalable aerosol particles with sizes less than 2.5 μm. It is necessary to arouse the stakeholders' cooperation in various industries across the supply chains to achieve SDGs (Zhang et al. 2019).

Supply chain emissions have a profound impact on the sustainability of economy and society. Fine particulate matter ($PM_{2.5}$) is a typical example of supply chain emissions, which is essentially a particulate matter emission. $PM_{2.5}$ emissions of supply chain mainly come from motorised-

CONTACT Zhenmin Li zhenmin.li@hfut.edu.cn

transportation means, including airplanes, trucks and ships that are dominating transportation forces of modern supply chain. $PM_{2.5}$ emissions have an adverse effect on public health (Yin, Pizzol, and Xu 2017). At present, many countries attach great importance to inhalable total suspended particulates that cannot be blocked by the human upper respiratory tract, especially inhalable aerosol particles with sizes less than 2.5 μm.

Particulate matter existing in our atmosphere, especially $PM_{2.5}$, can remain suspended in air for extended periods of time and has been adopted by the Chinese government as a representative pollutant for monitoring environmental atmospheric pollutants (Wu et al. 2020). The higher its concentration in air, the more severe the air pollution is. Although it is only a minimal component of atmospheric composition, $PM_{2.5}$ has a considerable effect on air visibility and quality. $PM_{2.5}$ features a smaller particle size, larger area and stronger activity, compared with bulkier atmospheric particles, and it is highly possible to assimilate harmful and toxic substances, such as microorganisms and heavy metals (Gu et al. 2018). In addition, $PM_{2.5}$ has a longer residence time in the atmosphere and a longer transportation distance, and has negative impacts on human health. Monitoring $PM_{2.5}$ concentrations is the first and foremost step toward the identification of $PM_{2.5}$ sources, and distribution patterns can provide insights into $PM_{2.5}$ pollution mitigation and prevention (Liu, Duan, and Chen 2020; Xu et al. 2020). Therefore, numerous countries worldwide have improved their capacity for environmental monitoring to provide crucial information-based support for related government departments (Sun et al. 2018).

With the rapid development in information technology, big data analytics have become an emerging data management method (Chaudhary et al. 2018; Patel and Desai 2019). Big data can effectively promote intelligent management of data, strengthen correlations among data, solve the problem of data redundancy and effectively promote the governing pattern (Lee, Kao, and Yang 2014). Big data is considered to be an important source of innovation and productivity (Santi et al. 2018). Additionally, this technique has become an important tool for analysing regional environmental quality and economic development and has been used extensively in resource management, environmental protection, and policy formulation (Benke and Benke 2018). Thus, big data analysis provides a methodological basis for scientific and objective understanding of the correlations between regional environments and economic development and offers corresponding data to improve economic-environmental zoning (Park et al. 2019). Although big data analytics are highly valuable in situational analysis of the environment, their application in integrated economic-environmental zoning has yet to be investigated (Verma, Agrawal, and Sharan 2016).

To more efficiently identify the patterns of air pollution emissions in the context of big data, an enhanced deep learning approach is presented. A prediction model of the $PM_{2.5}$ concentration is established in three stages, namely, data preparation, model training and inferencing. Finally, the practical data of $PM_{2.5}$ concentrations are used to verify the adaptability of the proposed method. The framework of air pollution emission based on the proposed boosting neural network approach can effectively predict the emission load of air pollution in terms of $PM_{2.5}$ concentrations. With this information, more effective measures can be formulated targeting air pollutant management policies, which help to bring about a coordinated development of the environment and economy for the society. We hope to provide an effective approach to forecast air pollution emissions. The proposed method can give a recommendation for related participants. It is similar to recommendations in the e-commerce field (Pan, Wu, and Olson 2017; Pan and Wu 2020).

The main contributions of this paper are summarised as follows: (1) we propose an integrated analytical framework for sustainability analysis of supply chain management through particulate matter emissions prediction. This analytical framework is able to perform sustainability analysis with accuracy and provide reliable insights for the legislation and implementation of green supply chain management. (2) We exploit a boosting ensemble meta-algorithm combined with a neural network method to obtain a prediction model for particulate matter emission prediction, which can better extract the features from the context of big data of air pollution emission and conduct predictions with high efficiency. (3) The effectiveness of the proposed framework is evaluated using 1598 effective particulate matter emission data ($PM_{2.5}$ and PM_{10}) and supply chain statistics from 2016 to 2019 in Beijing of China. (4) The sustainability analysis is of enlightening significance for policy-making targeting sustainable supply chain management to improve the regional environment and economy.

The remainder of this paper is organised as follows. Section 2 is a literature review. Section 3 describes an integrated analytical framework for sustainability analysis of supply chain management. Section 4 presents the experimental results and further analysis. Finally, we end the paper with conclusions in Section 5.

2. Literature review

There have been a number of researches on the carbon-sensitive supply chain. For instance, a problem of carbon-sensitive supply chain network with green procurement is introduced and analysed using a mixed-integer programming model, which incorporates carbon trading cost to reduce the overall amount of emissions (Abdallah, Diabat, and Simchi-Levi 2010). The combustion of fossil fuel and the production of heavy industry are deemed two major sources of direct ambient $PM_{2.5}$ emissions (Stevens and Boucher 2012). In recent years, there has been a drastic surge of anthropogenic primary $PM_{2.5}$ emissions in developing countries. This is owing to the increasing production activities, such as cement and power production caused by the fast development of economy and industrialisation (Lei et al. 2011). On the contrary, there has been an evident reduction of primary $PM_{2.5}$ emissions in developed countries. The main reasons include the transference of manufacturing activities and the promulgation of a set of national regulations regarding air quality (Meng et al. 2016). One significant finding is that the consumption of a commodity, its geographical primary production, and its final production are often fragmented in the present globalised world (Zhang and Lin 2018).

The variation trend of $PM_{2.5}$ in the context of big data can be used as a tool to classify patterns with similar emission characteristics into one category to abstract the common characteristics of pollution emission. There are extensive studies regarding air pollution monitoring technologies and the prediction of pollution emission load (Lung et al. 2020; Xu et al. 2020). $PM_{2.5}$ is affected by various factors, including meteorological factors such as barometric pressure, intensity of pressure, wind direction, wind speed, environmental factors, population factors and other air quality indices (Shen and Yao 2017). Therefore, it is possible to create information redundancy in high-dimensional $PM_{2.5}$ predictive modelling, which could cause unnecessary loading into the gating circular unit network if we cannot simplify the data analysis technology.

The study of $PM_{2.5}$ concentration trends has always been a research hotspot. Recently, deep learning methods, essentially complex nonlinear models, are extensively used in the prediction of air pollutant concentrations due to their usability and high performance. An extension to the conventional regression neural network was proposed to replace the point predictions with prediction intervals that meet a designated performance (Papadopoulos and Haralambous 2011). Song et al. (2014) achieved spatiotemporal $PM_{2.5}$ prediction using a spatial data-assisted incremental support vector regression method, with the data originated from 13 monitoring stations in New Zealand. A prediction model based on RBF neural network is presented for predicting $PM_{2.5}$ concentration, where a comparison with traditional BP network is further conducted (Zheng and Shang 2013). Li et al. (2017) presented an extension to short-term memory neural network to predict air pollutant concentration using historical air pollutant data, which integrated auxiliary data to improve performance. Zhou, Li, and Qiao (2017) presented a $PM_{2.5}$ prediction method using a recursive fuzzy neural network. Another machine learning algorithm has been proposed to deal with the spatial nonstationarity of the associations between $PM_{2.5}$ emissions and predictor variables (Zhan et al. 2017). Xiao et al. (2018) put forward a reliable $PM_{2.5}$ hindcast method using an ensemble machine learning technique. Chen et al. (2018) utilise a random forests model together with two traditional regression models to form a hybrid method to estimate ground-level $PM_{2.5}$

concentrations. Huang and Kuo (2018) combine a convolutional neural network and long short-term memory methods for implementing the $PM_{2.5}$ forecasting system.

3. Integrated analytical framework for sustainability analysis of supply chain management

3.1. Overall framework

The proposed integrated analytical framework for sustainability analysis of supply chain management is shown in Figure 1, which consists of four steps:

(1) The particulate matter emission data and supply chain statistics are collected and preprocessed. Conducting standardisation on the data can reduce the influence of varying dimensions on the analysis.
(2) The correlation analysis between air pollutant emissions and supply chain statistics is performed in order to evaluate the association and causality relation between air pollution and supply chain operations.
(3) Then, we perform particulate matter emissions prediction, using the boosting neural network approach that combines the neural network algorithm and boosting algorithm to improve the prediction performance of the model.
(4) We analyse the overall trend between air pollutant emissions and supply chain statistics after acquiring the particulate matter prediction results, serving as sustainability analysis and providing insights for green supply chain management policy formulation.

Since the data will also contain abnormal values, noise and missing values, it is necessary to clean and correct the data set to guarantee validity. Considering the running efficiency of the data mining algorithm, it is necessary to discretise the data and reduce the attributes to smooth the data and prepare for the prediction of particulate matter concentration.

In the intelligent interconnection environment, $PM_{2.5}$ concentration monitoring data present the characteristics of big data, namely, volume, velocity, diversity, and veracity (Aggarwal et al. 2013). In terms of information management, $PM_{2.5}$ big data is a type of information resource that can be used to reflect the degree of air pollution and has the characteristics of massiveness, heterogeneity, growth, complexity and repeatable mining. Profound data mining is needed to discover the information and knowledge that can aid in the comprehensive understanding of the emission of pollutants. The new high-dimensional, time-varying and large volume features have yielded better criteria for $PM_{2.5}$ classification, and new real-time processing methods are needed for concentration prediction. Thus, we propose a novel machine learning-based approach of $PM_{2.5}$ concentrations

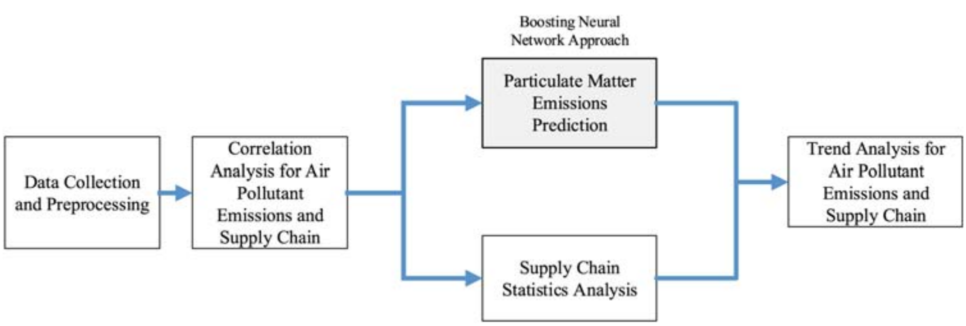


Figure 1. The proposed analytical framework for sustainability analysis of supply chain management.

using the boosting and neural network algorithms, and the framework of the proposed approach includes three principal stages: data preparation, boosting neural network and result referencing.

Daily $PM_{2.5}$ levels provided by air quality monitoring stations in various regions are collected via intelligent monitoring technology, and a database of these air pollution data is constructed. The characteristics of daily $PM_{2.5}$ concentration data are extracted from the preprocessed data, where other air pollution emissions, including PM_{10}, SO_2, CO, NO_2 and O_3, can also be used as input environmental data from which more features can be extracted. We choose the boosting algorithm integrated with the neural network approach for achieving the prediction model. The disadvantages of the neural network approach are mitigated by introducing the boosting ensemble meta-algorithm. Finally, the $PM_{2.5}$ concentration prediction results are obtained through model inferencing on test input data. The prediction results are visualised such that the variation trend of $PM_{2.5}$ concentration can be understood and evaluated intuitively. Therefore, a more effective strategy can be formulated targeting improving the air pollutant management, which can further benefit economic development and environmental protection through customised and accurate decision-making.

3.2. Boosting neural network approach for particulate matter emissions prediction

A simple neural network by itself has the problems of low classification accuracy and unstable prediction performance. In this study, an enhanced neural network is proposed to improve the prediction performance of the $PM_{2.5}$ concentration value model. This kind of enhanced neural network mainly combines the neural network algorithm and boosting algorithm to improve the prediction performance of the overall model, that is, using the boosting algorithm to train a certain number of individual subnets for the sample training set, and then weighted average the output results of each individual subnet to get the final prediction results.

We randomly divide the samples into a training sample set and test sample set according to the proportion of 7:3. For the input training sample set $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, we initialise a learner first $f_0(x) = \arg\min \sum_{i=1}^{n} L(y_i, \gamma)$, and then for $m = 1, \dots, M$, we set $M = 15$ here and repeat the following steps:

(1) Calculate the negative gradient $\overline{y_i} = -(\partial L(y_i, f_{m-1}(x_i)))/\partial f_{m-1}(x_i),\ i = 1, 2, \dots, n$
(2) By minimising the squared error, a basic learning machine $h_m(x)$ is used to fit $\bar{y}_i$ and here $w_m = \arg\min \sum_{i=1}^{n} [\bar{y}_i - h_m(x_i; w)]^2$
(3) Use line search to determine the step size ρ_m, where $\rho_m = \arg\min \sum_{i=1}^{n} L(y_i, f_{m-1}(x_i) + \rho h_m(x_i; w_m))$, to minimise L
(4) Obtain the learner of this iteration $f_m(x) = f_{m-1}(x) + \rho_m h_m(x; w_m)$. When the number of iteration reaches M, the final enhanced neural network prediction model $f_M(x)$ can be acquired, whereas $h_m(x) \approx -(\partial L(y, f_{m-1}(x)))/\partial f_{m-1}(x)$, $L(y_i, f(x_i)) = \frac{1}{2}[y_i - f(x_i)]^2$.

For each learner, a set of input data comes to the input layer, and then a set of data S is generated as the input of the hidden layer by the connection weight with the hidden layer, and then becomes $\theta(s_j)$ by the activation function of $\theta(s)$ of the hidden layer node, where s_j represents the output generated by the jth node of the hidden layer, which will generate the output layer by the connection weight between the hidden layer and the output layer. After the same processing as the hidden layer, the output y_j is generated at the output layer. Here, the specific formula of the activation function of $\theta(s)$ mentioned above is as follows:

$$\theta(s) = \frac{1}{1 + e^{-s}}.$$

For the prediction performance of the model, we mainly use the training sample set to learn the weight value to minimise the root-mean-square difference and the boosting algorithm. First, the loss function is defined as follows:

$$L(e) = \frac{1}{2}\sum_{j=0}^{k} e_j^2 = \sum_{j=0}^{k} (\bar{y}_j - y_j)^2.$$

4. Results

In this section, we present the experimental results in three aspects. After collecting and preprocessing the particulate matter emission data and supply chain statistics, we firstly perform the correlation analysis between supply chain statistics and particulate matter concentrations to reveal their associations. Secondly, we employ the boosting neural network approach combining the neural network algorithm and boosting algorithm to perform particulate matter concentration prediction. Finally, we analyse the overall sustainability between particulate matter concentrations and supply chain statistics after acquiring the particulate matter concentration results.

4.1. Experimental setup

In this work, SPSS modeller 18.0 and SPSS 25.0 are used as software tools for prediction and correlation analysis modelling. The parameter settings are presented in this subsection. (1) Particulate matter emission prediction: firstly, $PM_{2.5}$ is set as the target variable and other as the input by using the type node in SPSS modeller to complete the field role definition; secondly, the data are divided into a training set and test set according to the ratio of 7:3; then 10 MLP (multi-layer feedforward neural network) are selected as component models, and the boosting neural network prediction model is created by boosting; then SPSS is used to create the boosting neural network prediction model. The modeller automatically calculates the optimal number of hidden layers; in order to prevent the prediction model from overfitting, the overfitting prevention set is set to 30%. (2) Correlation analysis: SPSS was used to calculate the Pearson correlation coefficient among variables under the hypothesis test significance level of 0.05, and the correlation between the total volume of posts and telecommunications, total passenger transportation, total freight transportation, passenger turnover, freight turnover and $PM_{2.5}$ content was explored, respectively.

The supply chain statistics of Beijing from 1 January 2016 to 31 December 2019 are collected monthly from Beijing Municipal Bureau Statistics and Survey Office of the National Bureau of Statistics in Beijing (http://tjj.beijing.gov.cn/). Typical supply chain statistics including the volumes of post and telecommunication, passenger transport, freight transport, passenger turnover and freight turnover are collected and used for the experiments. The data of air pollutant emissions used for the experiments come from 1598 effective air quality data entries of the Beijing area from 1 January 2016 to 31 December 2019, obtained from China's online air quality monitoring platform (https://www.aqistudy.cn/historydata/). The platform is recognised by the state and its data are judged to be authoritative and authentic. Each data entry includes the concentration values of $PM_{2.5}$, PM_{10}, SO_2, Co, NO_2 and O_3 of Beijing on the same day.

4.2. Correlation analysis for air pollutant emissions and supply chain

We employ the Pearson correlation coefficient to measure the strength of the association between air pollutant emissions and supply chain data. The Pearson correlation coefficient formula between variables X and Y are given as

$$\rho X,\, Y = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}},$$

where n is the number of variable values. This formula returns a value between −1 and 1. A higher absolute value of the correlation coefficient represents a stronger correlation between target variables. The result of correlation analysis between air pollutant emissions and supply chain data in Beijing of China from 2016 to 2019 is shown in Table 1.

From Table 1, it is evident that the post and telecommunication volume in the year 2016 has a strong positive correlation with $PM_{2.5}$. The $PM_{2.5}$ emission in 2016 is the highest amongst these four years, suppressing transportation volumes other than posts and telecommunications. From the year 2017 to 2019, the volumes of passenger and freight transport exhibit a negative correlation with $PM_{2.5}$. Moreover, the passenger, freight and passenger turnover volumes in the year 2019 have a strong negative correlation with $PM_{2.5}$. It can be attributed to the reduction of air pollutant emissions continuously.

4.3. Particulate matter concentration prediction using boosting neural network

In order to acquire particulate matter concentration prediction results with high precision, we combine the boosting ensemble meta-algorithm and neural network method to train a set of individual subnets for the training set of samples, and then the output results of individual subnets are weighted average to get the final prediction results.

Table 2 summarises the detailed statistical distribution of each air pollutant. Using the $PM_{2.5}$ data of Beijing, the experiment is carried out, and the results of $PM_{2.5}$ prediction employing the boosting neural network method are shown in Figure 2. Each learner of a neural network used for integration is a three-layer MLP neural network structure. The number of hidden layer units is calculated automatically by the SPSS modeller according to the optimal performance.

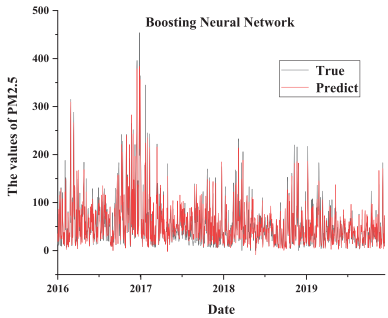
In order to further verify the effectiveness of the proposed method, we compare it with the current prominent machine learning methods, including classic neural network, linear regression, C&R and CHAID decision tree, and SVM; the corresponding experimental results are shown in Table 3. These methods are used extensively for data classification and prediction. Here, we use R-squared (R^2), mean squared error (MSE), root-mean-squared error ($RMSE$), and mean absolute error (MAE) as evaluation metrics.

Table 1. Correlation analysis for $PM_{2.5}$ emissions and supply chain in Beijing, China.

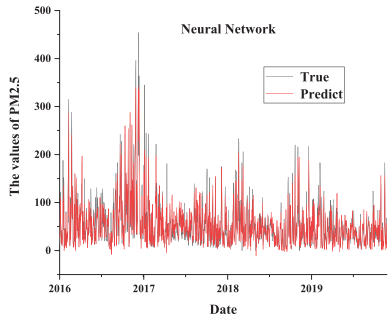
Year	Post and telecom	Passenger transport	Freight transport	Passenger turnover	Freight turnover
2016	0.627	−0.81	−0.003	−0.21	0.463
2017	0.607	−0.206	−0.560	−0.202	−0.435
2018	−0.141	−0.525	−0.334	−0.18	−0.251
2019	−0.293	−0.721	−0.703	−0.793	−0.277
2016–2019	−0.269	−0.293	−0.227	−0.563	−0.191

Table 2. Variable statistical distribution.

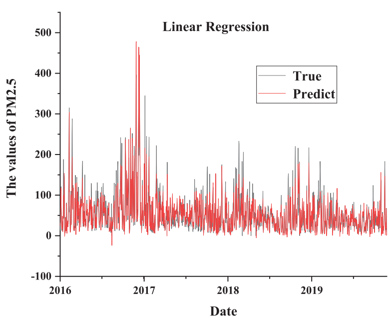
	Count	Minimum	Maximum	Mean	Standard deviation
$PM_{2.5}$	1598	0	454	54.11	49.955
PM_{10}	1598	0	512	78.87	56.450
SO_2	1598	0	84	6.67	7.123
CO	1598	0	8.0	0.891	0.6929
NO_2	1598	0	155	41.13	20.784
O_3_8h	1598	0	311	95.26	60.787



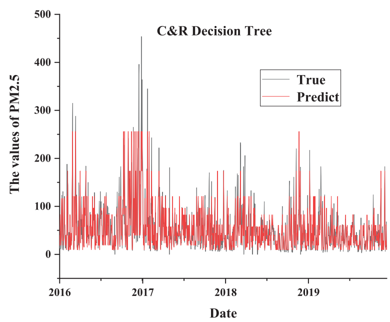
（a） Boosting Neural Network



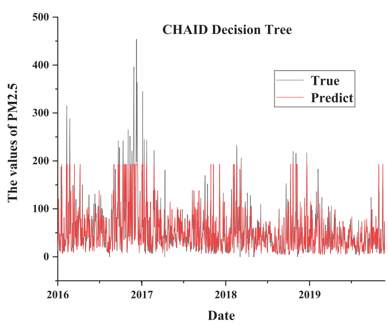
（b） Neural Network



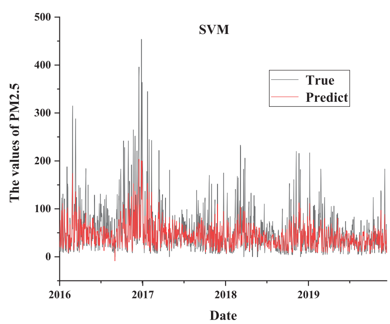
（c） Linear Regression



（d） C&R Decision Tree



（e） CHAID Decision Tree



（f） SVM

Figure 2. The result fitting curve of competing methods for $PM_{2.5}$ prediction.

Table 3. The performance comparison for $PM_{2.5}$ prediction.

Methods	R^2	MSE	$RMSE$	MAE
Boosting neural network	0.921	195.99	14.00	10.03
Neural network	0.894	265.43	16.29	11.13
Linear regression	0.861	346.92	18.63	12.53
C&R decision tree	0.845	385.60	19.64	12.36
CHAID decision tree	0.796	509.78	22.58	12.65
SVM	0.612	968.33	31.12	18.69

The formulae for calculating R^2, MSE, $RMSE$, MAE are given below:

$$MAE = \frac{1}{m}\sum_{i=1}^{m} |f_i - y_i|,$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m} (f_i - y_i)^2,$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m} (f_i - y_i)^2},$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (f_i - y_i)^2}{\sum_{i=1}^{m} (\overline{y_i} - y_i)^2}.$$

The R^2 is used to represent the quality of a fitting through the change of data. The MSE can be used to evaluate the change degree of data; it can be used to denote how accurate the prediction model can describe the experimental data. The $RMSE$ is used to measure the deviation between the observed value and the true value. The MAE can reflect the actual situation of prediction error.

In order to further verify the effectiveness of the proposed method, we compare it with the current prominent machine learning methods, including classic neural network, linear regression, C&R and CHAID decision tree, and SVM; the corresponding comparison results are shown in Table 4. Here, we use R-squared (R^2), mean squared error (MSE), root-mean-squared error ($RMSE$), and mean absolute error (MAE) as evaluation metrics.

We can observe from Table 3 and Figure 2 that the enhanced neural network prediction model achieved the best results, followed by the classic neural network model. Compared with the other four methods, our method achieved the largest R^2 value, meaning that our method can better capture the features between the $PM_{2.5}$ concentration and other variables. For MSE, $RMSE$ and MAE, our method obtained the smallest results, indicating that the best quality of $PM_{2.5}$ concentration prediction, which will bring greater help to the actual air pollutant control and management. In addition, it is evident that the enhanced neural network model is better than other methods in the closeness of fitting between the predicted value and the actual value obtained by each model. Therefore, the effectiveness of the proposed prediction model of $PM_{2.5}$ is demonstrated.

Table 4. The performance comparison for PM_{10} prediction.

Methods	R^2	MSE	$RMSE$	MAE
Boosting neural network	0.654	1091.42	33.04	20.62
Neural network	0.654	1092.16	33.05	20.62
Linear regression	0.630	1230.44	35.08	22.60
C&R decision tree	0.474	1744.16	41.76	26.75
CHAID decision tree	0.555	1523.79	39.04	24.97
SVM	0.242	1748.91	41.82	25.90

We also utilise the boosting neural network method to perform PM_{10} concentration prediction using air pollutant emission data of Beijing. The prediction results are shown in Figure 3, and the performance comparison is presented in Table 4. Similar to the results for $PM_{2.5}$ prediction, our method obtained the highest R^2 value, and lowest *MSE*, *RMSE* and *MAE* value, denoting the highest quality of PM_{10} prediction. Moreover, our method also achieved the best fitting closeness amongst all methods. Therefore, the effectiveness and applicability of our method on

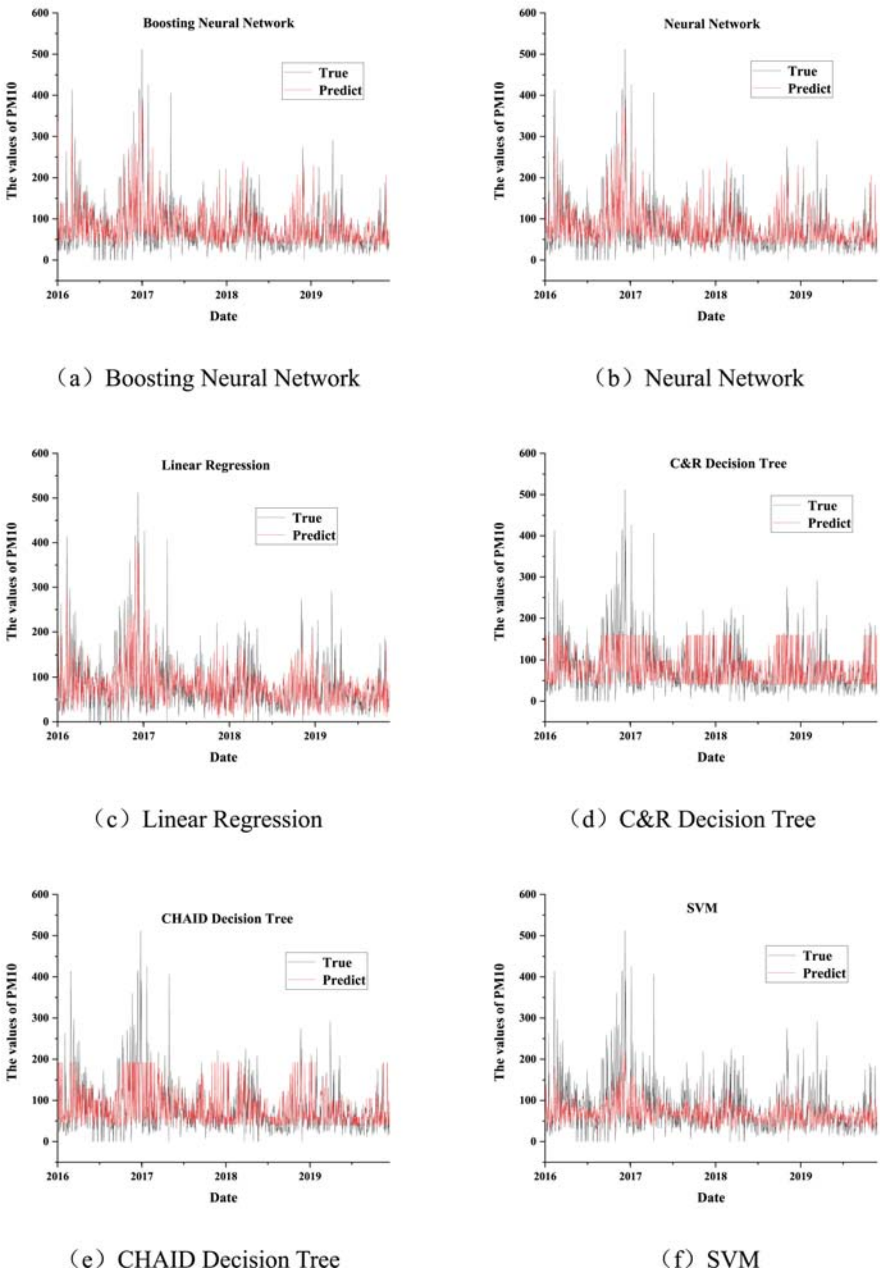


Figure 3. The result fitting curve of competing methods for PM_{10} prediction.

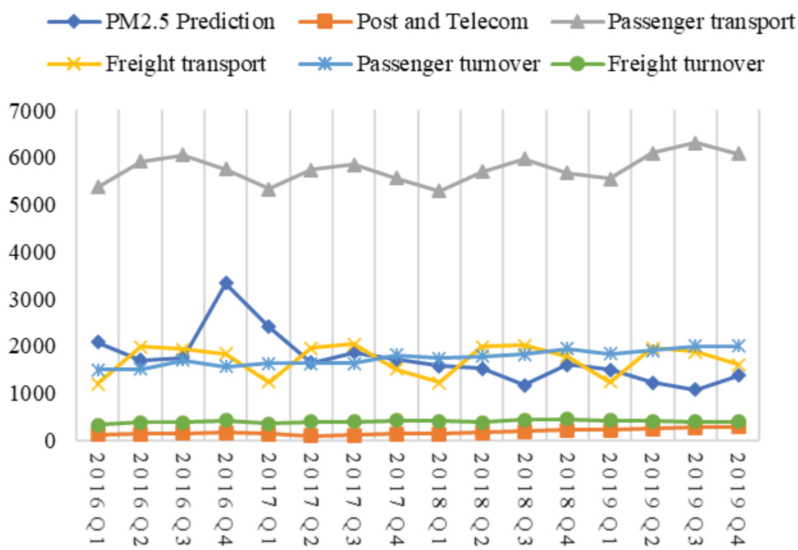


Figure 4. Sustainability analysis between particulate matter concentrations and supply chain statistics.

multiple particulate matter prediction are verified. Just like e-commerce, our proposed method can formulate recommendations for predicting air pollutant emissions (Pan et al. 2019; Pan and Wu 2019).

4.4. Sustainability analysis between particulate matter concentrations and supply chain statistics

We perform the sustainability analysis through overall trend evaluation between air pollutant emissions and supply chain statistics after acquiring the particulate matter prediction results. The quarterly volume changes of $PM_{2.5}$ concentration prediction, post and telecommunication, passenger transport, freight transport, passenger turnover and freight turnover are denoted as a line chart in Figure 4. It is evident that the $PM_{2.5}$ prediction is declining steadily from 2016 Q4 to 2019 Q3. Meanwhile, a slow upward trend is identified for the supply chain statistics including post and telecommunication, passenger transport, passenger turnover and freight turnover. Even though the freight transport volume fluctuates periodically, it is not affected by the decrement of $PM_{2.5}$ concentration. In sum, the sustainable increment of supply chain statistics has been realised from 2016 to 2019; in other words, the development of supply chain operations is not achieved at the cost of rising air pollutant emissions. The downward trend of $PM_{2.5}$ emission can be ascribed to the effectiveness of the legislations and regulations for air pollutant control in China.

The Air Law in China was issued firstly in 1987 and has undertaken several revisions. The latest major revision took place in 2015 and became effective on 1 January 2016 (Feng and Liao 2016). This revision stresses the air pollutant control induced by industrial production, coal burning, dust, motor vehicles, motor vessels, and agricultural activities. It establishes an explicit goal of enhancing the quality of air. It has detailed regulations for constraining greenhouse gases and atmospheric pollutants, specifically particulate matter, such as $PM_{2.5}$ and PM_{10} volatile organic compounds, nitrogen oxides, ammonia, and sulfur dioxide.

5. Conclusions

In this paper, an integrated analytical framework targeting sustainability analysis of supply chain management through particulate matter emissions prediction has been proposed. Specifically, we firstly collect and preprocess the particulate matter emission data and supply chain statistics in Beijing of China, then we conduct correlation analysis between air pollutant concentrations and supply chain statistics to uncover the intrinsic associations. Secondly, we combine the boosting ensemble meta-algorithm and neural network method for particulate matter concentration forecasting. The boosting algorithm is used to train a set of individual subnets to improve the efficiency of prediction, and afterwards, the output results of individual subnets are weighted for acquiring the final prediction results. Finally, we perform an analysis on sustainability between air pollutant emissions and supply chain statistics after obtaining the particulate matter prediction results.

This analytical framework is able to perform sustainability analysis with accuracy and provide reliable insights for the legislation and implementation of green supply chain management. The forecasting of particulate matter emissions conveys basis and justifications for policy-makers to enforce the environmental protection measures associated with air pollution and design efficient policies with the aim of improving the air pollutant management of the supply chain, which will further lead to an improved public health environment, especially in economically developed areas. Because this study mainly targets the prediction of $PM_{2.5}$ and PM_{10} concentrations, it cannot comprehensively reveal the geographical distribution of air pollution. There are various factors that affect air quality, and a regression model can be established to predict $PM_{2.5}$ and PM_{10} distributions more comprehensively and to determine the key influential factors, which would be important for formulating economic development strategies and implementing environmental protection mechanisms.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is fully supported by the Ministry of Education Engineering Research Center for Intelligent Decision-Making & Information System Technologies of China [No. PA2020GDGP0057].

References

Abdallah, T., A. Diabat, and D. Simchi-Levi. 2010. "A Carbon Sensitive Supply Chain Network Problem with Green Procurement." The 40th International Conference on Computers and Industrial Engineering: Soft Computing Techniques for Advanced Manufacturing and Service Systems (CIE40 2010), Awaji, July 25–28.

Aggarwal, Shankar G., Sudhanshu Kumar, Papiya Mandal, Bighnaraj Sarangi, Khem Singh, Jyoti Pokhariyal, Sumit K. Mishra, et al. 2013. "Traceability Issue in $PM_{2.5}$ and PM_{10} Measurements." *Mapan* 28 (3): 153–166.

Benke, Kurt, and Geza Benke. 2018. "Artificial Intelligence and Big Data in Public Health." *International Journal of Environmental Research and Public Health* 15 (12): 2796.

Chaudhary, Rajat, Gagangeet Singh Aujla, Neeraj Kumar, and Joel J.P.C. Rodrigues. 2018. "Optimized Big Data Management Across Multi-cloud Data Centers: Software-Defined-Network-Based Analysis." *IEEE Communications Magazine* 56 (2): 118–126.

Chen, Gongbo, Shanshan Li, Luke D. Knibbs, N. A. S. Hamm, Wei Cao, Tiantian Li, Jianping Guo, Hongyan Ren, Michael J. Abramson, and Yuming Guo. 2018. "A Machine Learning Method to Estimate $PM_{2.5}$ Concentrations Across China with Remote Sensing, Meteorological and Land Use Information." *Science of the Total Environment* 636: 52–60.

Feng, Lu, and Wenjie Liao. 2016. "Legislation, Plans, and Policies for Prevention and Control of Air Pollution in China: Achievements, Challenges, and Improvements." *Journal of Cleaner Production* 112: 1549–1558.

Gu, Xiaohe, Lei Wang, Wei Zhuang, and Lijian Han. 2018. "Reduction of Wheat Photosynthesis by Fine Particulate ($PM_{2.5}$) Pollution Over the North China Plain." *International Journal of Environmental Health Research* 28 (6): 635–641.

Huang, Chiou-Jye, and Ping-Huan Kuo. 2018. "A Deep CNN-LSTM Model for Particulate Matter ($PM_{2.5}$) Forecasting in Smart Cities." *Sensors* 18 (7): 2220.

Lee, Jay, Hung An Kao, and Shanhu Yang. 2014. "Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment." *Procedia CIRP* 16: 3–8.

Lei, Y., Q. A. Zhang, C. Nielsen, and K. B. He. 2011. "An Inventory of Primary Air Pollutants and CO_2 Emissions from Cement Production in China, 1990–2020." *Atmospheric Environment* 45: 147–154.

Li, Xiang, Ling Peng, Xiaojing Yao, Shaolong Cui, Yuan Hu, Chengzeng You, and Tianhe Chi. 2017. "Long Short-Term Memory Neural Network for Air Pollutant Concentration Predictions: Method Development and Evaluation." *Environmental Pollution* 231: 997–1004.

Liu, Hui, Zhu Duan, and Chao Chen. 2020. "A Hybrid Multi-Resolution Multi-objective Ensemble Model and Its Application for Forecasting of Daily $PM_{2.5}$ Concentrations." *Information Sciences* 516: 266–292.

Lung, Shih Chun Candice, Wen-Cheng Vincent Wang, Tzu-Yao Julia Wen, Chun-Hu Liu, and Shu-Chuan Hu. 2020. "A Versatile low-Cost Sensing Device for Assessing $PM_{2.5}$ Spatiotemporal Variation and Quantifying Source Contribution." *Science of the Total Environment* 716: 137145.

Meng, J., J. F. Liu, Y. Xu, D. B. Guan, Z. Liu, Y. Huang, and S. Tao. 2016. "Globalization and Pollution: Tele-connecting Local Primary $PM_{2.5}$ Emissions to Global Consumption." *Proceedings of the Royal Society A: Mathematical Physics and Engineering Series* 472. doi:10.1098/rspa.2016.0380.

Pan, Yuchen, and Desheng Wu. 2019. "Personalized Online-to-Offline (O2O) Service Recommendation Based on a Novel Frequent Service-Set Network." *IEEE Systems Journal* 13 (2): 1599–1607.

Pan, Yuchen, and Desheng Wu. 2020. "A Novel Recommendation Model for Online-to-Offline Service Based on the Customer Network and Service Location." *Journal of Management Information Systems* 37 (2): 563–593.

Pan, Yuchen, Desheng Wu, Cuicui Luo, and Alexandre Dolgui. 2019. "User Activity Measurement in Rating-Based Online-to-Offline (O2O) Service Recommendation." *Information Sciences* 479: 180–196.

Pan, Yuchen, Desheng Wu, and David L. Olson. 2017. "Online to Offline (O2O) Service Recommendation Method Based on Multi-Dimensional Similarity Measurement." *Decision Support Systems* 103: 1–8.

Papadopoulos, Harris, and Haris Haralambous. 2011. "Reliable Prediction Intervals with Regression Neural Networks." *Neural Networks* 24 (8): 842–851.

Park, Yong Shin, N. Muhammad Aslaam Mohamed Abdul Ghani, Fesseha Gebremikael, and Gokhan Egilmez. 2019. "Benchmarking Environmental Efficiency of Ports Using Data Mining and RDEA: The Case of a U.S. Container Ports." *International Journal of Logistics Research and Applications* 2 (22): 172–187.

Patel, Ashish Bhanabhai, and Tushar N. Desai. 2019. "A Systematic Review and Meta-analysis of Recent Developments in Sustainable Supply Chain Management." *International Journal of Logistics Research and Applications* 4 (22): 349–370.

Santi, Daniele, Elisa Magnani, Marco Michelangeli, Roberto Grassi, Barbara Vecchi, Gioia Pedroni, Laura Roli, et al. 2018. "Seasonal Variation of Semen Parameters Correlates with Environmental Temperature and Air Pollution: A Big Data Analysis Over 6 Years." *Environmental Pollution* 235: 806–813.

Shen, Yonglin, and Ling Yao. 2017. "$PM_{2.5}$, Population Exposure and Economic Effects in Urban Agglomerations of China Using Ground-Based Monitoring Data." *International Journal of Environmental Research and Public Health* 14 (7): 1–15.

Shi, Daqian, and Hongwei Yu. 2020. "Reevaluating the Subjective Welfare Loss of Air Pollution." *Journal of Cleaner Production* 257: 120445.

Song, Lei, Shaoning Pang, Ian Longley, Gustavo Olivares, and Abdolhossein Sarrafzadeh. 2014. "Spatio-Temporal $PM_{2.5}$ Prediction by Spatial Data Aided Incremental Support Vector Regression." International Joint Conference on Neural Networks (IJCNN), Beijing, July 6–11. IEEE.

Stevens, B., and O. Boucher. 2012. "Climate Science: The Aerosol Effect." *Nature* 490: 40–41.

Stieb, David M., Robin H. Shutt, Lisa M. Kauri, Sarah Mason-Renton, Li Chen, Mieczyslaw Szyszkowicz, Nina A. Dobbin, et al. 2019. "Associations Between Air Pollution and Cardio-Respiratory Physiological Measures in Older Adults Exercising Outdoors." *International Journal of Environmental Health Research*. doi:10.1080/09603123.2019.1699506.

Sun, Qinghua, Jia Zhuang, Yanjun Du, Dandan Xu, and Tiantian Li. 2018. "Design and Application of a Web-Based Real-Time Personal $PM_{2.5}$ Exposure Monitoring System." *Science of the Total Environment* 627: 852–859.

UN (United Nations). 2015. *Sustainable Development Goals*. Department of Public Information, United Nations. Accessed June 2020. http://www.un.org/sustainabledevelopment/sustainable-development-goals/.

Verma, Hanuman, R. K. Agrawal, and Aditi Sharan. 2016. "An Improved Intuitionistic Fuzzy C-means Clustering Algorithm Incorporating Local Information for Brain Image Segmentation." *Applied Soft Computing Journal* 46: 543–557.

Wu, Xueling, Ying Wang, Siyuan He, and Zhongfang Wu. 2020. "$PM_{2.5}/PM_{10}$ Ratio Prediction Based on a Long Short-term Memory Neural Network in Wuhan, China." *Geoscientific Model Development* 13 (3): 1499–1511.

Xiao, Qingyang, Howard H. Chang, Guannan Geng, and Yang Liu. 2018. "An Ensemble Machine-Learning Model to Predict Historical $PM_{2.5}$ Concentrations in China from Satellite Data." *Environmental Science & Technology* 52 (22): 13260–13269.

Xu, Zhaowei, Fang Lu, Wei Xu, Zhichao Wang, Yingxia Yang, Ce Xu, Pin Wu, and Xiaofeng Li. 2020. “Atmospheric Air Quality in Beijing Improved by Application of Air Source Heat Pump (ASHP) Systems.” *Journal of Cleaner Production* 257: 120582. doi:10.1016/j.jclepro.2020.120582.

Yin, Hao, Massimo Pizzol, and Linyu Xu. 2017. “External Costs of $PM_{2.5}$ Pollution in Beijing, China: Uncertainty Analysis of Multiple Health Impacts and Costs.” *Environmental Pollution* 226: 356–369.

Zhan, Yu, Yuzhou Luo, Xunfei Deng, Huajin Chen, Michael L. Grieneisen, Xueyou Shen, Lizhong Zhu, and Minghua Zhang. 2017. “Spatiotemporal Prediction of Continuous Daily $PM_{2.5}$ Concentrations Across China Using a Spatially Explicit Machine Learning Algorithm.” *Atmospheric Environment* 155: 129–139.

Zhang, Haoran, Ruixiong Li, Bin Chen, Huiming Lin, Qianru Zhang, Maodian Liu, Long Chen, and Xuejun Wang. 2019. “Evolution of the Life Cycle Primary $PM_{2.5}$ Emissions in Globalized Production Systems.” *Environment International* 131: 104996.

Zhang, Zengkai, and Jintai Lin. 2018. “From Production-Based to Consumption-Based Regional Carbon Inventories: Insight from Spatial Production Fragmentation.” *Applied Energy* 211: 549–567.

Zheng, Haiming, and Xiaoxiao Shang. 2013. “Study on Prediction of Atmospheric $PM_{2.5}$ Based on RBF Neural Network.” 4th international conference on Digital Manufacturing & Automation, Qingdao, June 29–30.

Zhou, Shanshan, Wenjing Li, and Junfei Qiao. 2017. “Prediction of $PM_{2.5}$ Concentration Based on Recurrent Fuzzy Neural Network.” 36th Chinese Control Conference (CCC), Dalian, July 26–28.