Front. Environ. Sci. Eng. 2022, 16(9): 119
https://doi.org/10.1007/s11783-022-1551-6

RESEARCH ARTICLE

Development of machine learning multi-city model for municipal solid waste generation prediction

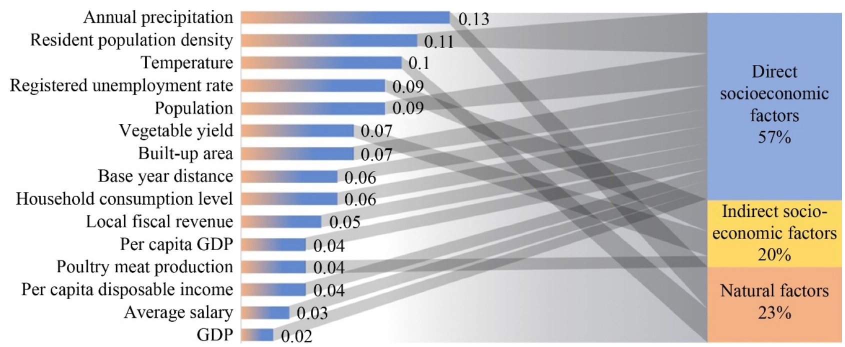
Wenjing Lu (✉), Weizhong Huo, Huwanbieke Gulina, Chao Pan

School of Environment, Tsinghua University, Beijing 100084, China

HIGHLIGHTS

- A database of municipal solid waste (MSW) generation in China was established.
- An accurate MSW generation prediction model (WGMod) was constructed.
- Key factors affecting MSW generation were identified.
- MSW trends generation in Beijing and Shenzhen in the near future are projected.

GRAPHIC ABSTRACT



ARTICLE INFO

Article history:
Received 5 October 2021
Revised 3 January 2022
Accepted 5 January 2022
Available online 26 February 2022

Keywords:
Municipal solid waste
Machine learning
Multi-cities
Gradient boost regression tree

ABSTRACT

Integrated management of municipal solid waste (MSW) is a major environmental challenge encountered by many countries. To support waste treatment/management and national macroeconomic policy development, it is essential to develop a prediction model. With this motivation, a database of MSW generation and feature variables covering 130 cities across China is constructed. Based on the database, advanced machine learning (gradient boost regression tree) algorithm is adopted to build the waste generation prediction model, i.e., WGMod. In the model development process, the main influencing factors on MSW generation are identified by weight analysis. The selected key influencing factors are annual precipitation, population density and annual mean temperature with the weights of 13%, 11% and 10%, respectively. The WGMod shows good performance with R^2 = 0.939. Model prediction on MSW generation in Beijing and Shenzhen indicates that waste generation in Beijing would increase gradually in the next 3–5 years, while that in Shenzhen would grow rapidly in the next 3 years. The difference between the two is predominately driven by the different trends of population growth.

1 Introduction

Ever increasing solid waste production has been threatening natural environment and human safety in recent years. With increasing urbanization worldwide, Municipal Solid Waste (MSW; the solid waste generated in the daily life of urban residents or in the process of serving daily life) has increased significantly (Iyamu et al., 2020). To tackle this challenge, various measures including reduction and resource recovery have been widely implemented to achieve better integrated solid waste management (Pires et al., 2011; Mukherjee et al., 2020). Notably, accurate prediction of MSW generation can greatly influence waste management system (Cherian and Jacob, 2012; Ghinea et al., 2016). It can directly affect the selection of subsequent treatment technology; the design of transport means, frequency and route; and the planning of treatment facilities. These would lay the foundation for the planning, implementation and optimization of the whole management system (Kumar et al., 2011). The prediction of MSW generation is a complex problem which requires a lot of historical data and various related factors (Kannangara et al., 2018). The commonly used

✉ Corresponding author
E-mail: luwenjing@tsinghua.edu.cn

prediction models include multi-variable linear regression model (Buenrostro et al., 2001; Mohammad Ali Abdoli, 2011; Azadi and Karimi-Jashni, 2016), time-series analysis model (Navarro-Esbrí et al., 2002; Marandi and Ghomi, 2016), gray system model (Huang et al., 1995; Chang and Pires, 2015) and system dynamics model (Eleyan et al., 2013). Although these traditional models are mature using simple methods (Abbasi and El Hanandeh, 2016), they usually choose a mathematical basic model in advance, which could limit the ability to truly reflect the characteristics of MSW (Abbasi et al., 2013).

Machine learning prediction models are increasingly used in solid waste management systems due to their high accuracy and ability to obtain new complex data and mine them in depth (Noori et al., 2010; Shahabi et al., 2012; Abbasi and El Hanandeh, 2016; Kontokosta et al., 2018). In addition, machine learning models can be broadly applied to short-, medium- and long-term predictions for MSW generation (Zade and Noori, 2008; Ali Abdoli et al., 2012; Abbasi et al., 2013). Machine learning algorithms such as artificial neural network (ANN) (Noori et al., 2010; Azadi and Karimi-Jashni, 2016), support vector machine (SVM) (Abbasi and El Hanandeh, 2016) and gradient boost regression tree (GBRT) (Johnson et al., 2017; Kontokosta et al., 2018) have been adopted for MSW generation forecasting. Relative to other algorithms, GBRT shows the following advantages. First, various types of data can be flexibly processed, including continuous values and discrete values. Second, in the case of relatively short tuning time, the accuracy of prediction can be relatively high. Third, the usage of some robust loss functions can be robust to outliers. The accuracy and practicability of model prediction are often conditioned by the selection and identification of feature variables (Ordóñez-Ponce et al., 2006;Adeogba et al., 2019).Whilea model simulation in Vietnam obtained an R^2 value > 0.96, that study merely used 63 detailed data sets to conduct machinelearningandgeographicdistribution(Nguyen et al., 2021). Leave-one-out or K-fold cross-validation can improve model accuracy especially for small data analysis. Cross-validation is a method of model selection, using part of the data set to test the model validity. However, only 12% of studies in a recent review on ANN studies have applied this method indicating its importance needs further attention (Xu et al., 2021). Extensive and comprehensive feature variables can further improve the model accuracy (Sun and Chungpaibulpatana, 2017). However, few studies have established the MSW generation model through multi-level feature variables (e.g., socioeconomic factors, natural conditions and internal conditions). Less than 10% of the published works on machine learning contained more than 1000 data in a report (Xu et al., 2021). In addition, in existing research, small scale data collection for most models aimed at the city level (Noori et al., 2010; Abbasi et al., 2014; Abbasi and El Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Johnson et al., 2017; Kannangara et al., 2018; Kontokosta et al., 2018; Wu et al., 2020), which limited the broad applicability of the model to a certain extent. Therefore, it is critical to develop a high-accuracy model based on large-scale data collection and wide range of influence variables that can be broadly applied to the prediction of MSW production.

To meet the needs of large-scale comprehensive treatment and realize the short-term MSW generation prediction, this study uses a wide range of data (country-wide city-based) from 130 cities across China and multi-level feature variables (e.g., socioeconomic factors, natural conditions and internal conditions) to establish a machine learning multi-city model of MSW generation with high accuracy. It is applied to analyze and explore the waste management modes of two typical large cities in China.

2 Materials and methods

2.1 Process for waste generation model development

To develop, test and optimize the waste generation model, the following steps are taken (Fig. 1):

1) Construction of database of MSW generation and feature variables (socioeconomic and natural) through large-scale city-level data collection and analysis across China.

2) Application of the Machine learning Gradient Boost Regression Tree (GBRT) algorithm to develop the prediction model upon the database established.

3) Identification of the main influential factors for MSW generation based on weight.

4) Optimization of the algorithm through iterations.

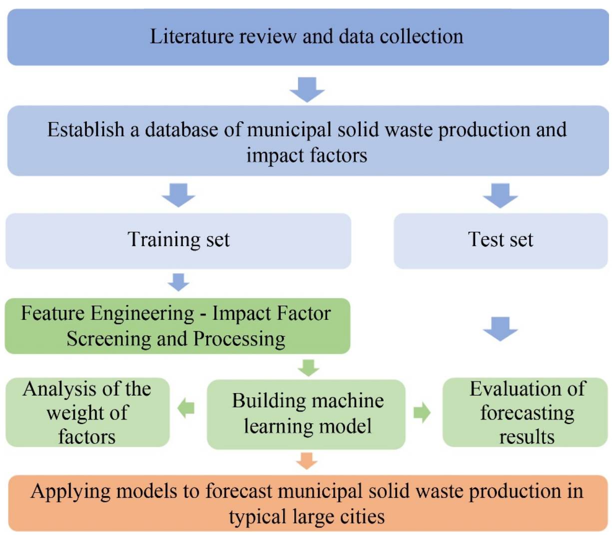


Fig. 1 Model construction method.

5) Model training through multiple epochs to obtain high testing accuracy.

2.2 Data screening and database construction

Here the data were screened and collected from various sources, e.g., urban statistical yearbook retrieval for socioeconomic data; National Meteorological Information Center for meteorological data; and literature for internal data. A sum of 2250 data sets (including MSW generation, economic, sociological and natural conditions) for 312 cities in China was acquired. To perform data pretreatment, the following processes are conducted: 1) Missing value processing (delete data sets that are missing a certain characteristic value) and 2) abnormal value processing (delete data sets with large errors caused by external factors, e.g., operational errors). After data screening, a total of 1012 complete data sets involving 130 cities were obtained. A complete data set includes 27 feature variables of a city, and the actual amount of MSW generation in 2000–2017. The 27 feature variables (see Table 1) consist of three categories: direct socioeconomic factors, indirect socioeconomic factors and natural factors. Direct socioeconomic factors refer to population, urban construction level, economy and other factors directly related to the amount of MSW generation (Al-Salem et al., 2018). In the database of this study, there are 10 direct socioeconomic factors, e.g., population, built-up area and GDP. Indirect socioeconomic factors refer to the socioeconomic factors with some impact on the MSW generation (yet are not directly related); eight variables are included, e.g., registered unemployment rate and vegetable yield (Bashir and Goswami, 2016; Zoroufchi Benis et al., 2019). The database also includes nine natural factors (e.g., geographical location and temperature) which can affect the MSW generation to certain extend (Kontokosta et al., 2018).

Table 1 Feature variables related to MSW generation

Factor type	Feature variables	Unit
Internal factors	Population	Ten thousand people
	Built-up area	Square kilometer
	Resident population density	Person per square kilometer
	Gross domestic product (GDP)	Billion
	Per capita GDP	Yuan
	Local fiscal revenue	Billion
	Household consumption level	–
	Per capita disposable income	Yuan
	Average salary	Yuan
	Base year distance	Years
Socioeconomic factors	Urban residents' vegetable consumption expenditure	Yuan
	Registered unemployment rate	%
	Vegetable yield	Ton
	Poultry meat production	Ton
	Land utilization	%
	Cleaning street area	Square kilometer
	Higher education ratio	%
	Tourism activity income	Ten thousand yuan
Natural factors	Geographic location (South/North)	–
	Geographic location (East/West)	–
	The annual average temperature	Celsius
	Annual precipitation	Millimeter
	Annual average wind speed	Meter/second
	Average pressure	Hapa
	Windy days	Day
	Rainy days	Day
	Climate type	–

The finalized database covers 31 provincial-level administrative regions with 130 cities (the corresponding locations are marked on the map in Fig. 2(a)). Due to the differences in economic development of different regions in China, the obtained data are mostly from the more developed regions, especially the south-east coastal area. Since the number of cities in the western region is less (than that in the eastern region), hence there are less data from the western region. Nevertheless, the data set in general covers a wide range and has a good overall representation in the country.

Cities in China are classified into five categories (Table 2): super-mega city (more than 10 million people), mega city (5 to 10 million people), large city (1 to 5 million people), medium city (0.5 to 1 million people) and small city (< 0.5 million people). According to the classification criteria (above), the data of 130 cities are statistically analyzed. From the city size perspective, it shows a centralized trend in geographical distribution (Fig. 2(b)). Nevertheless, all five categories of cities are covered because each category constitutes at least 10% of all cities. Since the database covers a wide area and contains large differences, it can ensure that the machine learning models trained using the database have the advantages of strong generality and applicability to various categories of cities.

2.3 Construction of machine learning model

2.3.1 Training and testing data sets

The machine learning model of MSW generation is developed based on the aforementioned database. By taking the annual MSW generation of the city as the dependent variable y_i and 27 feature variables as independent variables x_i, their relationship can be expressed as Eq. (1):

$$x_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(27)}), \tag{1}$$

where i is the i_{th} group of data; x_i^k is the k_{th} feature value

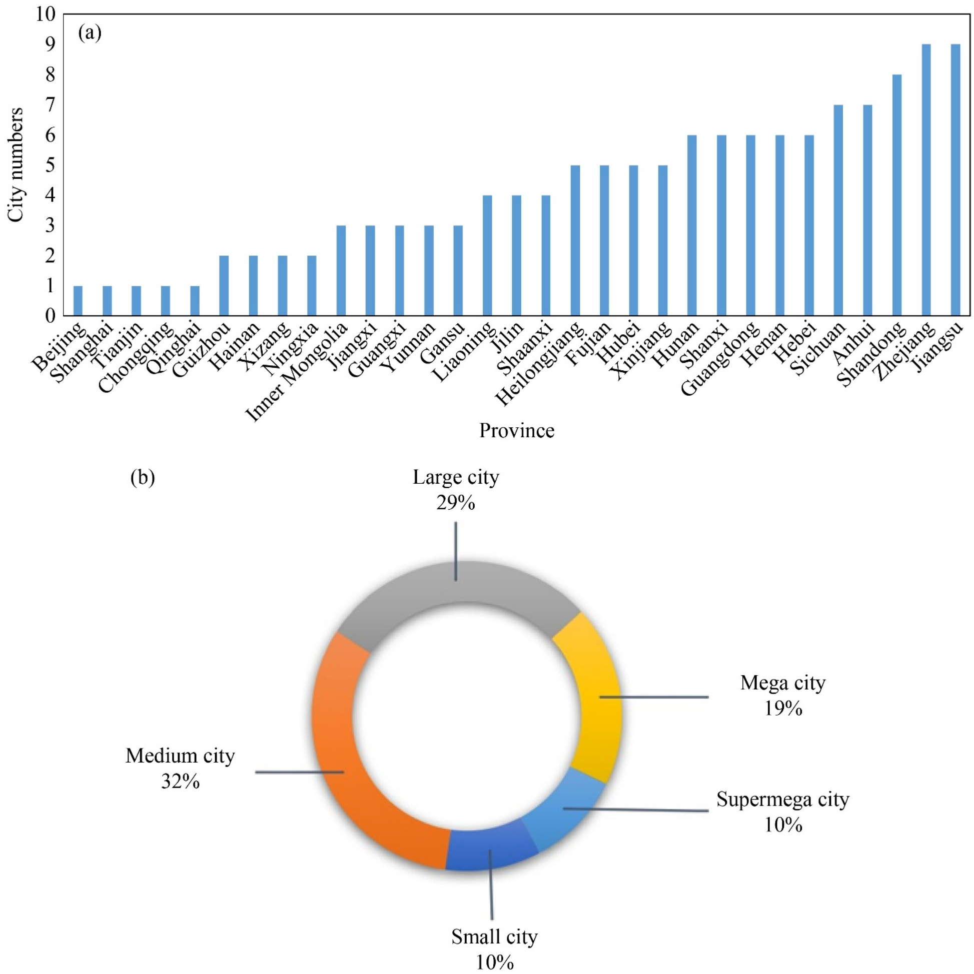


Fig. 2 Coverage of the city-level data in terms of city composition and geographical distribution (a) and proportion of cities with various sizes (b).

Table 2 Prediction results of the WGMod for different categories of cities

City category	Population	Proportion	WGMod-R^2
Mega-cities	> 5 million	29%	0.893
Large cities	1–5 million	29%	0.943
Medium cities	500000–1 million	32%	0.961
Small cities	< 500000	10%	0.958

of x_i.

In model development process, the data are randomly divided into 80:20, i.e., the model randomly selects 80% of the entire data (1012) as training set and its remaining 20% as testing set.

2.3.2 Model parameters determination

The maximum number of iterations is the most important parameter of decision tree. When the number of iterations is too small, underfitting is likely to occur; when it is too large, overfitting will occur (Miller et al., 2016). According to the empirical method, the best number of iterations in the gradient boost regression tree is 100 when the sample size is < 2000. Besides, a higher iterations number can achieve a better fitting result when the sample size is larger than this number. Since there are 1012 data sets in this study, the number for iterations is set to 100.

Another important parameter of the model is the maximum depth of the decision tree. Since the number of independent variables (i.e., feature variables) in this study is 27 and the data size is 1012, none of them are too large. Thus, the depth of a single decision tree is not limited for this mode.

In addition, the maximum number of feature variables used by each tree is also a parameter to be considered during model construction. For Gradient Boost Regression Tree algorithm, in general, if the number of feature variables is not large (e.g., less than 50 in this model), the maximum number of features should not be limited. At the same time, since the factors that affect the amount of MSW generated in real world are much more complicated, the maximum number of feature variables in the model should not be limited in order to fully account for the impact of various factors on MSW generation.

2.3.3 Iterative framework for the model

As aforementioned, the total 1012 sets of data of machine learning model are divided into training and testing sets at the ratio of 80:20. Hence, the total number of samples in the training set is 817.

The input equation of the training set in the model is (Eq. (2)):

$$T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_{817}, y_{817})\}, x_i \subseteq R^n, y_i \in R, \quad (2)$$

where x is the feature variables and $x_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(27)})$; y_i is the annual MSW generation of the city; R is the set of real numbers; and n is the number of features.

When the prediction result of a single decision tree in iteration m is $f_m(x)$, the framework of the model is as follows:

1) Initialize $f_0(x) = 0$.

2) Each iteration of decision tree the residual is calculated by Eq. (3):

$$r_{mi} = y_i - f_{m-1}(x_i), \quad (3)$$

where r_{mi} is the residual; $m \in [1, 2, \ldots, 100]$ is the iteration; and $i \in [1, 2, \ldots, 817]$ is the training data.

3) Fit the residual r_{mi} and feature variables in training set to obtain a regression tree, which is noted as $T(x; \Theta_m)$, where T is the abstract function of the regression tree; x is the feature variable in the training set; and Θ_m indicates the characters of regression tree in iteration m.

4) Update the regression model to $f_m(x) = f_{m-1}(x) + T(x; \Theta_m)$.

5) Repeat steps 2), 3) and 4) until the iteration reaches to 100.

6) Set up the final gradient boost regression tree after 100 iterations (Eq. (4)):

$$f_M(x) = \sum_{m=1}^{100} T(x; \Theta_m). \quad (4)$$

2.4 Analysis of the model results

In this study, Spyder 3.3.1 is used as the development environment of Python (III) to construct the MSW generation model. Its built-in algorithm is used to evaluate the model accuracy and analyze the weights of the feature variables. The coefficient of determination R^2 is used to evaluate the testing accuracy of the model, and the weights of different feature variables are used to evaluate their importance to MSW generation.

2.5 Model prediction for typical large cities in China

Based on the influencing factors, the model is applied to predict the MSW generation of typical cities in China. Two cities, Beijing and Shenzhen, are selected as the simulation targets. To predict the MSW generation in the two cities, the data set of 27 feature variables of the WGMod were first gathered from the city statistical yearbooks of Shenzhen (period: 2006–2017) and Beijing (period: 2000–2017). They are inserted into the mathematical regression model to predict the values of these variables in 2018–2022 (see detailed data in Supplementary Information). The new variables are subsequently inserted into the established model for predicting MSW generation.

3 Results and discussion

3.1 Model training and iteration process

The MSW generation prediction model, i.e., WGMod, is an iterative optimization process based on the prediction results of each decision tree. As shown in Fig. 3, the total error of the WGMod on the training and testing data sets decreases with the number of iterations. When the number of iterations reaches 100, the total error value stabilizes. The result confirms that the reasonable limit is 100 iterations and there is no overfitting in the model prediction process (Miller et al., 2016). The blue line shows the model iteration process of the training data sets. After model training, the total error (i.e., between the predicted and true values) is less than 5. In the testing set (red line), the total error is less than 10 after 100 iterations, indicating the continuous model improvement through iterative training. Basic model parameters were shown Table S1.

3.2 Identification of main influential factors of municipal solid waste generation

WGMod is used for identifying the main influencing factors of MSW generation through weight analysis. By accumulating and analyzing the feature weights of the 100 decision trees during iterative process in the model, the weight distribution bar can be obtained (refer to Fig. 4). In the process of analyzing the decisive role of the factors, the variables with weight below 0.01 are excluded. Among the 27 feature variables, the natural variable of “annual precipitation” has the highest weight

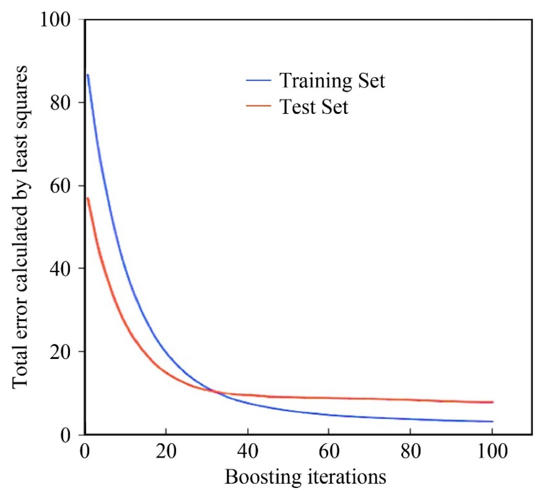


Fig. 3 Variance of prediction results of the WGMod as affected by the number of iterations.

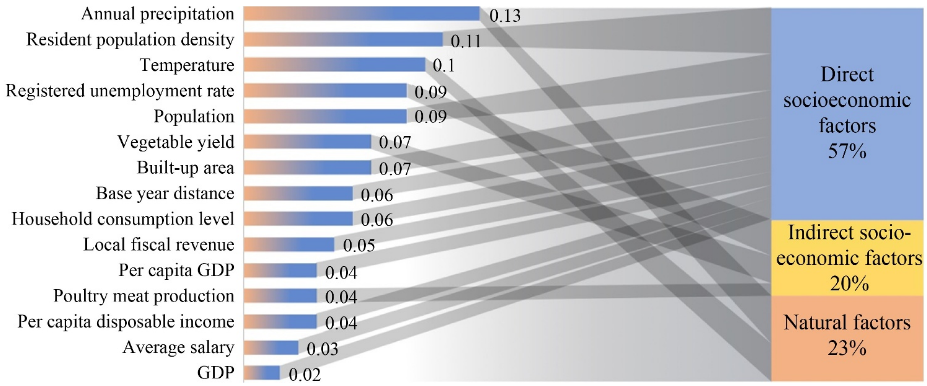


Fig. 4 Influencing factors of MSW generation identified by weight analysis of the WGMod.

of 13%. The annual precipitation is one of the most important factors affecting the amount of municipal waste because of two key reasons. One reason is that precipitation affects the mass weight of MSW by increasing their moisture content because there were fewer covers, shadings in storage and collection devices in the past. Another reason is that the annual precipitation (as one of the climate variables) affects the output of vegetables and vegetable consumption in the region, hence influences the composition of MSW. Therefore, the mass weight of MSW increases with the input of vegetable waste. The internal variable of population density ranks the second, with a weight of 11%. Obviously, when the per capita output of MSW is relatively stable, the population density is positively related to the total output of MSW (Khajevand and Tehrani, 2019). The third influencing factor is the natural variable of annual mean temperature with a weight of 10%. People's consumption habits can vary greatly in different regions (related to different climate type, temperature or season), because the differences in composition and quantity of products can result in different amounts of MSW (Purcell and Magette, 2009; Mohammad Ali Abdoli, 2011). Moreover, the output of garden waste can vary significantly with temperature variation (Boldrin and Christensen, 2010).

In addition, some variables (i.e., population, registered unemployment rate, built-up area, vegetable yield, household consumption level, GDP per capita, local fiscal revenue) all contribute to the model, and their weights are all greater than 5%. The remaining feature variables contribute 17% of the weight in total, indicating their relatively small influence on the amount of MSW generation. Three categories of feature variables (i.e., direct socioeconomic variables, indirect socioeconomic variables and natural variables) in the WGMod all have considerable weights (i.e., 57%, 20% and 23%, respectively) (Fig. 4). The weight analysis suggests that the population and built-up area driven by urbanization are the most important socio-economic influencing factors of MSW generation. While rainfall and temperature are the major natural influencing factors, indicating some effect of urban climate on MSW generation. Therefore, the key factors with weight > 1% are selected for predicting MSW generation.

3.3 Accuracy analysis of the WGMod

Following previous studies, a total of 1012 sets of data were divided randomly into training set and testing set at the ratio of 80:20 for WGMod processing (Park et al., 2018; Roh et al., 2018; Xu et al., 2021). Therefore, 195 sets of data were used each time for testing the accuracy of WGMod. The coefficient of determination R^2 (Fig. 5) is used for judging the accuracy. The result shows the coefficient of determination R^2 of 0.9390, reflecting good performance of the WGMod. Table 3 presents the R^2 values of different models of MSW generation developed using machine learning algorithms (Abbasi and El Hanandeh, 2016; Kontokosta et al., 2018;Wu et al., 2020).Adeogba et al. developed a model for food and garden waste prediction in UK using GBRT method (Adeogba et al., 2019). Although a high-accuracy machine learning model (R^2 > 0.96) was developed for major cities in Vietnam, the applied data sets were relatively small (Nguyen et al., 2021). In the current study, more feature variables and larger coverage area (multi-cities and cities of different sizes) were considered. Besides, few multi-city models

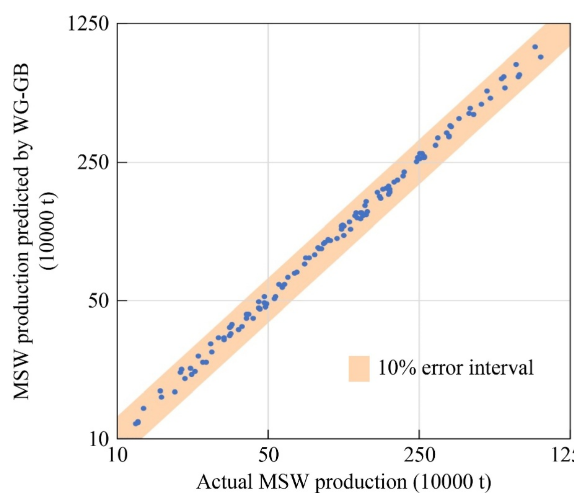


Fig. 5 MSW production predicted by WGMod versus actual MSW production.

Table 3 Comparison of different MSW machine learning prediction models

MSW prediction model	Algorithm	Data Sources	Prediction accuracy R^2	Reference
WGMod	Gradient boosting regression trees	130 cities in China	0.94	This study
LR model	Random forest	Czech Republic	0.77	Rosecky et al.,2021
M5Tree	Model tree	Kahrizak dumpsite, Iran	0.85	Alidoust et al., 2021
DNN	Deep neural network	Vietnam	0.91	Nguyen et al., 2021
Forecasting model	Gradient boosting regression trees	327 UK local authorities	0.65	Adeogba et al., 2019
GBRT model	Gradient boosting regression trees	New York, USA	0.87	Kontokosta et al., 2018
MSW-Census (Decision Trees)	Classification and Regression Tree	Ontario, Canada	0.54	Kannangara et al., 2018
MSW-Census (Neural Networks)	Single hidden layer feed forward neural network	Ontario, Canada	0.72	Kannangara et al., 2018
GBRT model	Gradient boosting regression trees	New York, USA	0.88	Johnson et al., 2017
ANFIS	Adaptive neuro-fuzzy inference system	Logan City Council region in Queensland, Australia	0.98	Abbasi and El Hanandeh, 2016
kNN	K-nearest neighbors	Logan City Council region in Queensland, Australia	0.51	Abbasi and El Hanandeh, 2016
ANN model	Artificial neural network	Fars province, Iran	0.67–0.86	Azadi and Karimi-Jashni, 2016
GT/PCA/-ANN models	Artificial neural networks	Mashhad, Iran	0.73–0.80	Noori et al., 2010

have been developed for developing countries due to the slow establishment of the solid waste management system in developing countries and limited data availability. Limited data from China's lower-level administrative units could be a weak point of WGMod for predicting MSW in those areas.

The model managed to cover all city types with MSW generation ranges from 100000 t/a to 12.5 million t/a. Due to the geographical distribution of the original data set, the testing data set randomly selected by the model randomly were from cities located in the eastern part of China. It demonstrated good dispersion and representatives with small absolute deviation of the testing data set (0%–10%). As aforementioned, the 130 cities in the database were divided into four categories of cities, and the prediction accuracy of each category was calculated (Table 2). The result shows that for mega-cities (with a population size of above 5 million), model accuracy using the testing set (R^2 = 0.893) was the lowest. This might be related to some complex factors (e.g., huge amount of mobile population; variation in economic index; and faster implementation of new policies).

Except the city size, the model accuracy has no significant correlation with the category of the city and the geographical distribution evidenced by higher accuracy and a certain degree of randomness. The error distribution of the cases ranged 0.01–0.08, of which error less than 0.07 accounted for 84% of the cases, error less than 0.03 accounted for 29% of the cases (Figs. 6(a) and 6(b)).

3.4 Prediction of MSW generation for typical cities of China

Shenzhen and Beijing are important cities with strategic positions of socioeconomic development in the southern and northern China, respectively. The WGMod is applied to analyze the trend of MSW generation in these two cities.

The model result shows that the amount of MSW generated in Shenzhen would continue to rise in the next three years, and the projected annual growth rate is 4.5%–7.1% (Fig. 7(a)). The amount of MSW would reach 7.88 million tons by 2022. In early 2018, Shenzhen was selected by the Chinese government as a demonstration city for "Zero-waste". Therefore, in addition to the corresponding expansion of the scale of treatment facilities, especially the biochemical treatment system for perishable organic waste, minimization strategy of MSW should be promoted.

For Beijing, the amount of MSW generated in the next three years would also continue to increase (Fig. 7(b)). By 2022, the amount of MSW generated in Beijing would reach 11.22 million tons. However, the growth rate of Beijing's MSW generation would experience a significant downward trend. Relative to 2021, it would drop by 2.7% in 2022, and the five-year average growth rate is 4%, which means that the growth rate of MSW might gradually decrease. According to the trend of population decline in Beijing in 2017 and 2018, the apparent decrease in the growth rate of MSW generation might be related to the decrease in mobile population which caused a corresponding decline in the total population in the next

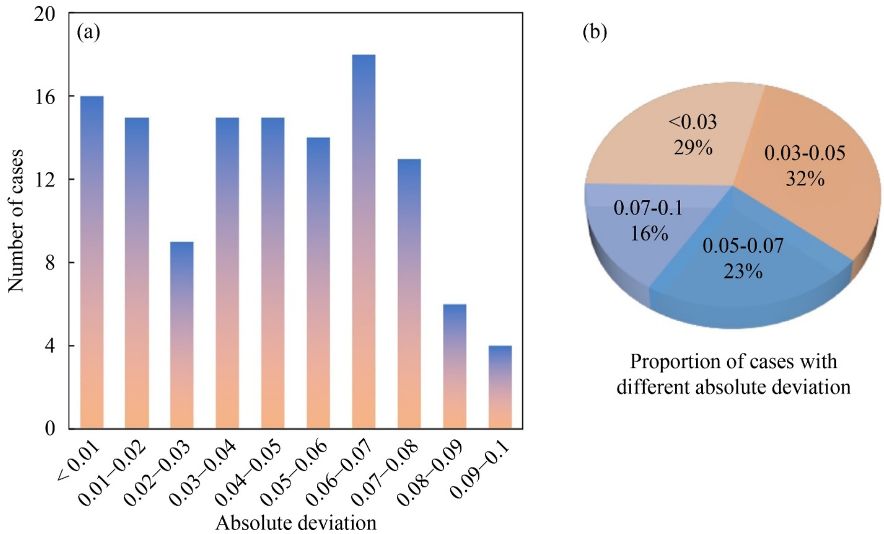


Fig. 6 Prediction accuracy and deviation of data set. Case distribution of different absolute deviation (a); proportion of cases with different absolute deviation (b).

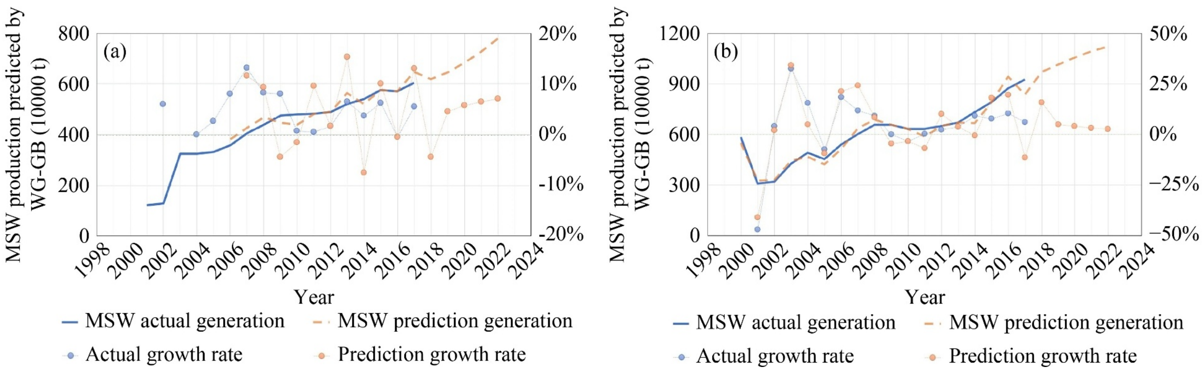


Fig. 7 MSW generation predicted by the WGMod for Shenzhen (a) and Beijing (b).

few years. Should the economic and sociological data maintain a stable trend, the MSW generation in Beijing might enter a plateau phase in the next 5–10 years.

4 Conclusions

Through literature review and data extraction, a database of MSW generation and feature variables with 1012 data sets covering 130 cities across China was established. The waste generation predicting model, i.e., WGMod, developed using gradient boost regression tree algorithms performed reasonably well with the coefficient of determination (R^2) of 0.939. Annual precipitation (13%), population density (11%) and average temperature (10%) were identified as the key influencing factors of MSW generation. The model was applied to predict the MSW generation in Beijing and Shenzhen. The results suggested that waste production in Beijing would grow slowly in the next 3–5 years, while that in Shenzhen would grow rapidly with an annual growth rate of 7.1% by 2022. This study provided scientific methods and basic data for a multi-city model development for MSW generation. Specifically, the WGMod is suitable for predicting MSW generation in China because it was developed based on Chinese database. Since the impact of variables on MSW generation could be different in different countries, it is necessary to further investigate/refine the influencing factors that meet the actual conditions of the target country.

Acknowledgements This work was financially supported by the National Key R&D Program of China (Nos. 2018YFD1100600 and 2018YFC1902900). We deeply appreciate our friends for correction of language and advice to this paper.

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11783-022-1551-6 and is accessible for authorized users.

References

Abbasi M, Abduli M A, Omidvar B, Baghvand A (2013). Forecasting municipal solid waste generation by hybrid support vector machine and partial least square model. International Journal of Environmental Research, 7(1): 27–38

Abbasi M, Abduli M A, Omidvar B, Baghvand A (2014). Results uncertainty of support vector machine and hybrid of wavelet transform-support vector machine models for solid waste generation forecasting. Environmental Progress & Sustainable Energy, 33(1): 220–228

Abbasi M, El Hanandeh A (2016). Forecasting municipal solid waste generation using artificial intelligence modelling approaches. Waste Management (New York, N.Y.), 56: 13–22

Adeogba E, Barty P, O'Dwyer E, Guo M (2019). Waste-to-resource transformation: Gradient boosting modeling for organic fraction municipal solid waste projection. ACS Sustainable Chemistry & Engineering, 7(12): 10460–10466

Al-Salem S M, Al-Nasser A, Al-Dhafeeri A T (2018). Multi-variable regression analysis for the solid waste generation in the State of Kuwait. Process Safety and Environmental Protection, 119: 172–180

Ali Abdoli M, Falah Nezhad M, Salehi Sede R, Behboudian S (2012). Longterm forecasting of solid waste generation by the artificial neural networks. Environmental Progress & Sustainable Energy, 31(4): 628–636

Alidoust P, Keramati M, Hamidian P, Amlashi A T, Gharehveran M M, Behnood A (2021). Prediction of the shear modulus of municipal solid waste (MSW): An application of machine learning techniques. Journal of Cleaner Production, 303: 127053

Azadi S, Karimi-Jashni A (2016). Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars Province, Iran. Waste Management (New York, N.Y.), 48: 14–23

Bashir S, Goswami S (2016). Tourism induced challenges in municipal solid waste management in hill towns: Case of Pahalgam. Procedia Environmental Sciences, 35: 77–89

Boldrin A, Christensen T H (2010). Seasonal generation and composition of garden waste in Aarhus (Denmark). Waste Management (New York, N.Y.), 30(4): 551–557

Buenrostro O, Bocco G, Vence J (2001). Forecasting generation of urban solid waste in developing countries: A case study in Mexico. Journal of the Air & Waste Management Association, 51(1): 86–93

Chang N, Pires A (2015). Grey Systems Theory for Solid Waste Management. Piscataway: IEEE Press

Eleyan D, Al-Khatib I A, Garfield J (2013). System dynamics model for hospital waste characterization and generation in developing countries. Waste Management & Research, 31(10): 986–995

Ghinea C, Drăgoi E N, Comăniţă E D, Gavrilescu M, Câmpean T, Curteanu S, Gavrilescu M (2016). Forecasting municipal solid waste generation using prognostic tools and regression analysis. Journal of Environmental Management, 182: 80–93

Huang G H, Baetz B W, Patry G G (1995). Grey quadratic programming and its application to municipal solid waste management planning under uncertainty. Engineering Optimization, 23(3): 201–223

Iyamu H O, Anda M, Ho G (2020). A review of municipal solid waste management in the BRIC and high-income countries: A thematic framework for low-income countries. Habitat International, 95: 102097

Cherian J, Jacob J (2012). Management models of municipal solid waste: A review focusing on socio economic factors. International Journal of Finance & Economics, 4: 131–139

Johnson N E, Ianiuk O, Cazap D, Liu L, Starobin D, Dobler G, Ghandehari M (2017). Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City. Waste Management (New York, N.Y.), 62: 3–11

Kannangara M, Dua R, Ahmadi L, Bensebaa F (2018). Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches. Waste Management (New York, N.Y.), 74: 3–15

Khajevand N, Tehrani R (2019). Impact of population change and unemployment rate on Philadelphia's waste disposal. Waste Management (New York, N.Y.), 100: 278–286

Kontokosta C E, Hong B, Johnson N E, Starobin D (2018). Using machine learning and small area estimation to predict building-level municipal solid waste generation in cities. Computers, Environment and Urban Systems, 70: 151–162

Kumar J S, Subbaiah K V, Rao P V V P (2011). Prediction of municipal solid waste with RBF net work: A case study of Eluru, A. P, India. International Journal of Innovation, Management and Technology, 2(3): 238−243

Marandi F, Ghomi S M T F (2016). Time series forecasting and analysis of municipal solid waste generation in Tehran city. In: Proceedings of the 12th International Conference on Industrial Engineering (ICIE). Tehran, Iran: ICIE2016, 14–18

Miller P J, Lubke G H, McArtor D B, Bergeman C S (2016). Finding structure in data using multivariate tree boosting. Psychological Methods, 21(4): 583–602

Mohammad Ali Abdoli M F (2011). Multivariate econometric Approach for solid waste generation modeling: Impact of climate factors. Environmental Engineering Science, 28(9): 627–633

Mukherjee C, Denney J, Mbonimpa E G, Slagley J, Bhowmik R (2020). A review on municipal solid waste-to-energy trends in the USA. Renewable & Sustainable Energy Reviews, 119: 109512

Navarro-Esbrí J, Diamadopoulos E, Ginestar D (2002). Time series analysis and forecasting techniques for municipal solid waste management. Resources, Conservation and Recycling, 35(3): 201–214

Nguyen X C, Nguyen T T H, La D D, Kumar G, Rene E R, Nguyen D D, Chang S W, Chung W J, Nguyen X H, Nguyen V K (2021). Development of machine learning-based models to forecast solid waste generation in residential areas: A case study from Vietnam. Resources, Conservation and Recycling, 167: 105381

Noori R, Karbassi A, Salman Sabahi M (2010). Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. Journal of Environmental Management, 91(3): 767–771

Ordóñez-Ponce E, Samarasinghe S, Torgerson L (2006). Artificial neural networks for assessing waste generation factors and forecasting waste generation: a case study of Chile. Journal of Solid Waste Technology Management, 32: 167–184

Park Y, Kim M, Pachepsky Y, Choi S H, Cho J G, Jeon J, Cho K H (2018). Development of a nowcasting system using machine learning approaches to predict fecal contamination levels at recreational beaches in Korea. Journal of Environmental Quality, 47(5): 1094–1102

Pires A, Martinho G, Chang N B (2011). Solid waste management in European countries: A review of systems analysis techniques.

Journal of Environmental Management, 92(4): 1033–1050

Purcell M, Magette W L (2009). Prediction of household and commercial BMW generation according to socio-economic and other factors for the Dublin region. Waste Management (New York, N.Y.), 29(4): 1237–1250

Roh S B, Park S B, Oh S K, Park E K, Choi W Z (2018). Development of intelligent sorting system realized with the aid of laser-induced breakdown spectroscopy and hybrid preprocessing algorithm-based radial basis function neural networks for recycling black plastic wastes. Journal of Material Cycles and Waste Management, 20(4): 1934–1949

Rosecky M, Somplak R, Slavik J, Kalina J, Bulkova G, Bednar J (2021). Predictive modelling as a tool for effective municipal waste management policy at different territorial levels. Journal of Environmental Management, 291: 112584

Shahabi H, Khezri S, Ahmad B B, Zabihi H (2012). Application of artificial neural network in prediction of municipal solid waste generation (case study: Saqqez city in Kurdistan Province). World Applied Sciences Journal, 20(2): 336–343

Sun N, Chungpaibulpatana S (2017). Development of an appropriate model for forecasting municipal solid waste generation in Bangkok. Energy Procedia, 138: 907–912

Wu F, Niu D, Dai S, Wu B (2020). New insights into regional differences of the predictions of municipal solid waste generation rates using artificial neural networks. Waste Management (New York, N.Y.), 107: 182–190

Xu A, Chang H, Xu Y, Li R, Li X, Zhao Y (2021). Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. Waste Management (New York, N.Y.), 124: 385–402

Zade J G, Noori R (2008). Prediction of municipal solid waste generation by use of artificial neural network: A case study of Mashhad. International Journal of Environmental Research, 2(1): 13–22

Zoroufchi Benis K, Safaiyan A, Farajzadeh D, Khalili Nadji F, Shakerkhatibi M, Harati H, Safari G H, Sarbazan M H (2019). Municipal solid waste characterization and household waste behaviors in a megacity in the northwest of Iran. International Journal of Environmental Science and Technology, 16(8): 4863–4872