Long short-term memory neural network and improved particle swarm optimization–based modeling and scenario analysis for municipal solid waste generation in Shanghai, China

Deyun Wang1 · Ying-an Yuan1 · Yawen Ben1 · Hongyuan Luo1,2 · Haixiang Guo1

Received: 19 January 2022 / Accepted: 20 April 2022 / Published online: 14 May 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Accurate estimations of municipal solid waste (MSW) generation are vital to effective MSW management systems. While various single-point estimation approaches have been developed, the non-linearity and multiple site-specific influencing factors associated with MSW management systems make it challenging to forecast MSW generation quantities precisely. To address these concerns, this study developed a two-stage modeling and scenario analysis procedure for MSW generation and taking Shanghai as a test case demonstrated its viability. In the first stage, nine influencing factors were selected, and a hybrid novel forecasting model based on a long short-term memory neural network and an improved particle swarm optimization (IPSO-LSTM) was proposed for the forecasting of the MSW generation quantities, after which actual Shanghai data from 1980 to 2019 were used to test the performance. In the second stage, the future influencing variable values in different scenarios were predicted using an improved grey model, after which the predicted Shanghai MSW generation quantities from 2025 to 2035 were evaluated under various scenarios. It was found that (1) the proposed IPSO-LSTM had higher accuracy than the benchmark models; (2) the MSW generation quantities are expected to respectively increase to 9.971, 9.684, and 9.090 million tons by 2025 and 11.402, 11.285, and 10.240 by 2035 under the low, benchmark, and high scenarios; and (3) the MSW generation differences between the high and medium scenarios were decreasing.

Keywords Municipal solid waste · Forecasting · Long short-term memory · Improved particle swarm optimization · Improved GM (1,1) · Scenario analysis

Abbreviations

AI Artificial intelligence
ANFIS Adaptive neuro-fuzzy inference system
ARIMA Autoregressive integrated moving average
CNN Convolutional neural network
CPI Consumer price index
CUPS Comprehensive Urban Planning of Shanghai (2017–2035)
ESM Exponential smoothing model
GDP Gross domestic product
GM Grey modeling
IoT Internet of things
LSTM Long short-term memory
MAE Mean absolute error
MAPE Mean absolute percentage error
MLR Multiple linear regression
MSW Municipal solid waste
NOT Number of overseas tourists
PRP Permanent resident population
PSO Particle swarm optimization
RNN Recurrent neural network
RMSE Root mean square error
RPI Retail price index
TI Tertiary industry
TIC Theil inequality coefficient
TIFA Total investment in fixed assets
VMD Variational mode decomposition
WtE Waste to energy
WtM Waste to material
UPCCE Urban per capita consumption expenditure
UPCDI Urban per capita disposable income

Responsible Editor: Eyup Dogan

Deyun Wang
wang.deyun@hotmail.com

1 School of Economics and Management, China University of Geosciences, Wuhan 430074, China
2 Nanomedicine Lab, Université de Bourgogne Franche-Comté, UTBM, 90010 Belfort, France
Introduction

Due to China’s rapid urbanization and economic development over the past few decades, municipal solid waste (MSW) generation quantities have increased significantly, making China the second largest global MSW producer behind the USA. MSW is heterogeneous comprising biodegradable waste, medical waste, electrical and electronic waste, and construction waste. Some waste, such as electronic products (phones, televisions, and computers), are hazardous, do not degrade quickly, and therefore can have negative impacts on the environment and people’s health. However, if MSW is disposed of correctly, it can be converted into valuable raw materials, such as gardening and agricultural fertilizer, energy, solid fuel, and higher-value byproducts (Liu et al., 2018). Because of the lack of adequate MSW infrastructures in many places and especially in developing areas, MSW is often poorly managed, resulting in poor environmental protection and resource utilization. Adequate and sustainable waste management infrastructure planning depends on the ability to reliably estimate future MSW generation (Kolekar et al., 2016), which requires a consideration of the expected demographic, social, and economic factor changes; therefore, accurate MSW forecasting is complex and challenging (Nguyen et al., 2021a). MSW quantitative prediction methods can be divided into three main categories: causal relationship forecasting methods, time-series forecasting methods, and hybrid forecasting methods.

Causal relationship MSW generation forecasting models

Causal relationship forecasting methods based on correlation prediction principles require an approximate expression of the functional relationships between the influencing factors and MSW generation (Ghinea et al., 2016). Therefore, accurate MSW generation predictions depend upon the selection of pertinent social and economic development influencing factors in the study area. Based on the quantitative interdependence relationships between the variables, the causal relationship forecasting MSW prediction methods can be divided into either linear forecasting models or nonlinear forecasting models.

Linear causal relationship models for MSW generation forecasting

Multiple linear regression (MLR), which uses polynomials to reflect the functional relationships between various influencing factors and MSW generation, has been one of the most commonly used linear regression (LR) analysis models. Araiza-Aguilar et al. (2020) used MLR to explore the impact of economic, social, and demographic explanatory variables on MSW generation and found that the model had excellent prediction performances. However, the LR method prediction accuracy highly depends on the rationality and comprehensiveness of the selected indicators, and as the MSW generation system is complex, multivariable, and nonlinear, linear functional relationships are generally unable to adequately describe the relationships between the independent and dependent variables. The highly nonlinear characteristics of the variables also limit forecasting accuracy and any extensive application of LR models.

Nonlinear causal relationship models for MSW generation forecasting

As artificial intelligence (AI) models, such as the artificial neural network (ANN), the adaptive neuro-fuzzy inference system (ANFIS), decision trees (DT), support vector machines (SVM), and fuzzy logic (FL), can mimic the human reasoning process, they have been extensively used for highly nonlinear MSW generation systems (Abdallah et al., 2020). ANNs, such as the radial basis function (RBF), multi-layer perceptron (MLP), and backward propagation (BP), have been the most commonly used AI models, followed by SVM. Despite their excellent prediction accuracy, ANNs have rigid input sequence constraints and disregard long-term impacts in their sequential data analysis (Niu et al., 2021). To overcome these shortcomings, due to their effectiveness in extracting input series characteristics, deep learning approaches, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention algorithms, have been used for modeling and forecasting in recent years (Sun and Ge, 2021). While RNNs have been typically employed to describe long-term span series nonlinear dynamic processes, they still suffer from gradient explosions and vanishing, which means that long sequential modeling has significant challenges (Yuan et al., 2020). Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997) established a long short-term memory (LSTM) neural network to address these issues, which employed the RNN to efficiently memorize the historical input sequence evolution law and then during the training procedure used the distinctive LSTM three-gate structure and cell state to evaluate the temporal span influence when processing input sequences in batches to determine the duration of the impacts (Solano Meza et al., 2019). Significant air quality (Yan et al., 2021), wind power (Shahid et al., 2021), and solar energy consumption (Nguyen et al., 2021b) research has applied the LSTM neural network to causal relationship analysis; however, few studies have employed it for MSW generation forecasting.
Time-series methods for MSW generation forecasting

Time-series models, which employ historical data statistical analysis on the dependent variable to determine the variation patterns (Chang and Lin, 1997), have been successfully applied to forecast the seasonal and periodic variations in MSW generation (Navarro-Esbrí et al., 2002). There are three most often used time-series methods: conventional statistical methods, decomposition-ensemble-based approaches, and AI-based methods. ARIMA, a conventional time-series analytical approach, has been widely employed in MSW generation prediction because of its sophisticated mathematical theory and its capacity to depict data reproducibility (Chao, 2008). The decomposition-ensemble-based approach, which was inspired by the “divide and conquer” concept, involves time-series decomposition, modal prediction, and integrated prediction (Yu et al., 2008) and has been successfully implemented to predict e-waste. For example, Wang et al. (2021) used variational modal decomposition (VMD) to decompose time-series data, exponential smoothing model (ESM) to forecast each time–frequency signal, and grey modeling (GM) to integrate each time–frequency signal, with the experimental findings demonstrating that the VMD-ESM-GM technique was able to accurately predict future e-waste fluctuation trends. In contrast to the highly nonlinear MSW generation systems described by AI regression models, the AI time-series models are able to describe the fluctuation trends in pure time series and anticipate the future. Sunayana and Kumar (2021) employed a nonlinear auto-regressive (NAR) neural network to forecast short-term monthly MSW generation in Nagpur, India, until 2023. However, if there is no or poor MSW generation historical data, AI-based methods may not be effective.

Hybrid models for MSW generation forecasting

In general, hybrid models deal with integrated problems or perform diverse tasks (Abdallah et al., 2020). There are three main hybrid model types: combined causal relationship forecasting methods, combined diverse time series approaches, and cross mixing approaches. Xu et al. (2013) developed a hybrid method that integrated the seasonal auto-regressive integrated moving average (SARIMA) and a GM(1,1) model to forecast month-scale, medium-term, and long-term MSW output, and then validated its performance using data from Xiamen, China. Chen and Chang (2000) proposed a hybrid forecasting approach based on grey theory, fuzzy theory, and system dynamics (GFM) and validated the model’s efficiency using solid waste production 1985 to 1998 time-series data from Tainan, China, finding that the GFM method prediction accuracy was substantially higher than the single GM model. Taking 1990 to 2018 Shanghai MSW data as an example, Lin et al. (2021) selected 25 generation factors and then developed a hybrid forecasting method based on a one-dimensional CNN, LSTM, and attention mechanism to depict the highly nonlinear MSW generation system. However, these hybrid approaches have limitations. For example, because time series methods have poor fitting capacity on sharp change points, it is difficult to anticipate the MSW generation from the formation mechanism. Although the causal relationship forecasting approach can avoid these issues, its application to forecasting future MSW generation is restricted as it does not know the value of each influencing factor in the future, and even though current hybrid MSW generation forecasting studies combine different causal relationship and time-series methods, few studies have combined the two to address the implementation deficiencies and identify the complex MSW production causalities.

Main contribution of this study

Taking Shanghai as an example, which is a highly developed Chinese city with a large resident population, this study proposes a hybrid forecasting model to accurately predict the city’s MSW generation from 2025 to 2035. First, an improved GM(1,1) is employed to effectively predict the various future influencing factors, after which an IPSO is used to optimize the LSTM hyper-parameters and obtain the prediction model that has the best fitting effect. Finally, a scenario analysis approach is adopted to evaluate the 2025 to 2035 MSW generation and provide realistic proposals for a viable environmental protection governance system in Shanghai.

Therefore, the main contributions and innovations are as follows: (1) a novel combined adjustment strategy is proposed to improve the global search capacity and convergence efficiency of the conventional PSO algorithm; (2) a novel hybrid model (IPSO-LSTM) based on an improved PSO algorithm and the LSTM neural network is proposed to ensure precise MSW generation forecasting; (3) a novel improved GM(1,1), the IGM(1,1), is established for the influencing factor time series forecasting from 2025 to 2035; and (4) a scenario analysis method based on IPSO-LSTM and IGM(1,1) is proposed that makes up for the shortcomings in single time series forecasting or causal relationship forecasting methods, accounts for the diverse future MSW generation growth trends, and pays attention to the factors and coordination relationship analyses crucial to the system’s development.

The remainder of this paper is structured as follows: “Material and methods” details the proposed hybrid forecasting model construction process, “Proposed model performance test” compares and verifies the effectiveness of IPSO-LSTM and IGM(1,1). “Shanghai MSW generation scenario analyses” adopts the IGM(1,1) to predict the
influencing factor values under different development prospects and accurately forecast the MSW generation in each IPSO-LSTM model scenario, and “Conclusions” provides some concluding remarks and valuable extensions to the current study.

Material and methods

Description of the study area

As Shanghai is a leading demonstration city for China’s waste classification policy, its trash categorization and recycling management policies are needed to strengthen the city’s image and soft power. However, because of Shanghai’s increasing business activities, there has been a significant escalation in consumption, which in turn has led to a significant rise in MSW generation. Figure 1 shows the MSW generation on the Chinese mainland in 2019, from which it can be seen that Shanghai’s waste generation is comparatively high. The Comprehensive Urban Planning of Shanghai (2017–2035) (CUPS) explicitly stated that by 2035, Shanghai is expected to be a socialist metropolis with an international impact. Therefore, accurate MSW generation forecasting in Shanghai over the medium to long term is vital for the planning and construction of related infrastructures, such as waste collection, transportation, and incineration; the establishment of scientific, comprehensive, environmental protection management systems; and ecological city sustainability.

Data sources and influencing factor choice

The data for this study were collected from Shanghai statistical yearbooks from the official Shanghai Bureau of Statistics website (http://tjj.sh.gov.cn). MSW generation is impacted by multiple elements, with most MSW studies having focused on social, economic, and demographic influencing variables. Therefore, based on previous MSW generation studies (Kannangara et al., 2018; Liu et al., 2021; Nguyen et al., 2021a) (Fig. 2), nine factors were selected: permanent resident population (PRP); the number of overseas tourists (NOT); gross domestic product (GDP); tertiary industry (TI); the consumer price index (CPI), the retail price index (RPI); urban per capita consumption expenditure (UPCCE); urban per capita disposable income (UPCDI); and the total investment in fixed assets (TIFA). The historical data and change trends for these nine factors are shown in Fig. 3. A correlation test was separately conducted for the MSW generation, the results for which are shown in Table 1, which shows that the Pearson’s correlation coefficients between these nine influencing factors and MSW were greater than 0.8, which indicated that all were significantly associated with MSW and could be used to forecast MSW generation. Considering there exists a high correlation relationship between all explanatory variables (except RPI), which may cause the “Multicollinearity” phenomenon and may negatively influence the model’s robustness and stability, this research has adopted L2 regularization (Deng et al. 2022), also known

Fig. 1 The thermal distribution of MSW production on the Chinese mainland in 2019
as ridge regression, to settle this problem based on the problem scale and data characteristics.

Figure 4 shows the MSW quantities in Shanghai from 1980 to 2019, from which it can be seen that there was a general growth tendency from 1.31 million tons in 1980 to 10.38 million tons in 2019. There was an unexceptional upsurge in 2000 and 2001, with no significant change in the relevant influencing factors in the same period, which was primarily the result of a government campaign that permitted the kitchen trash from the catering sector to be discharged into the MSW (Xiao et al., 2020). Therefore, to ensure prediction accuracy, the Shanghai MSW data collected from the National Bureau of Statistics (http://www.stats.gov.cn) in 2000 and 2001 were replaced.
The LSTM neural network is a peculiar RNN that performs better than a conventional RNN in describing historical input sequence evolutionary rules. As shown in Fig. 5, the key to the LSTM is the information transmission path throughout the cell state and the information addition and removal process. The cell state (long-term influence) is determined from the multiplication of the preceding and current cell states, which is calculated as

\[ C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \]  

where \( C_t \) is the new state, \( f_t \) is the forget gate output, \( C_{t-1} \) is the previous cell state, \( i_t \) is the input gate output, and \( \tilde{C}_t \) is the candidate state. The long-term influence is modulated by LSTM using a unique three-layer gate (forget gate, input gate, and output gate) structure. The forget gate determines the duration of the long-term effect and uses a Sigmoid
activation function to project an input sequence between 0 and 1 to select the information to be retained from the previous cell state. The following is a description of the forget gate functions:

\[
f_t = \sigma(W_f \times [X_t, h_{t-1}] + b_f)
\]  

where \( f_t \) is the forget gate output, \( \sigma \) is the Sigmoid activation function, \( W_f \) and \( b_f \) donate the linear relationship coefficient and bias to the forget gate, \( X_t \) is the input sequence, and \( h_{t-1} \) is the previous hidden state.

The input gate adds the immediate state to the long-term effect by selecting the information to be updated using the Sigmoid function, after which the input sequence is compressed between \(-1\) and \(1\) using the \( \tanh \) function. The input gate mathematical expressions are shown as follows:

\[
i_t = \sigma(W_i \times [X_t, h_{t-1}] + b_i)
\]  

\[
\tilde{C}_t = \tanh(W_c \times [X_t, h_{t-1}] + b_c)
\]  

where \( W_i \) and \( b_i \) in formula (3) respectively are the input gate linear relationship coefficient and bias, and in formula (4), \( \tanh \) donates the hyperbolic tangent activation function, and \( W_c \) and \( b_c \) are the candidate state linear relationship coefficient and bias.

The output gate is responsible for determining whether the long-term effect should be considered and controls the output of the current cell state with the aid of the sigmoid and \( \tanh \) activation functions, as follows:

\[
o_t = \sigma(W_o \times [X_t, h_{t-1}] + b_o)
\]  

\[
h_t = o_t \cdot \tanh(C_t)
\]  

where \( o_t \) in formula (5) is the outcome output gate, and \( W_o \) and \( b_o \) are the output gate linear relationship coefficient and bias, and the \( h_t \) in formula (6) is the new hidden state and \( C_t \) is a new state.

Therefore, LSTM overcomes the short-term memory influence because of the “three-tier gate” structural design, which allows for the information from the former time steps to be reflected in the current cell state.

### Improved particle swarm optimization

Particle swarm optimization (PSO), which was developed by Eberhart and Kennedy in 1995 based on bird hunting behavior, is a population-based heuristic method for addressing optimization problems (Marini and Walczak, 2015). The PSO algorithm seeks an optimal solution by collaborating and sharing information with the individuals in the group, that is, it simulates a flock of birds by manufacturing a massless particle that only has speed and position attributes. The PSO is initialized as a set of random particles, with each particle separately searching for the optimal solution in the solution space. In each iteration, based on the current personal best (\( p_{best} \)) and current global best (\( g_{best} \)) values, the particle calculates its speed for the next moment, after which it updates its position. Assuming there are \( j \) particles
in a given region, the \((k + 1)\) velocity vector update formula and position vector update formula for the \(i\)th particle are as follows:

\[
\begin{align*}
    v_i^{k+1} &= wv_i^k + c_1r_1 (p_{\text{best}}^k - p_i^k) + c_2r_2 (g_{\text{best}}^k - p_i^k) \\
    p_i^{k+1} &= p_i^k + v_i^{k+1}
\end{align*}
\] (7)

where parameter \(w\) denotes the inertia weight, which illustrates the impact of the previous iteration’s speed on the present iteration; the higher the value, the better the global optimization ability and vice versa; \(c_1\) and \(c_2\) are the individual learning and population learning coefficients that respectively indicate the local optimization and global optimization abilities, and \(r_1\) and \(r_2\) are random digits between 0 and 1.

The traditional PSO algorithm has a premature convergence problem, that is, each particle moves to the current optimal local value while searching, which gradually diminishes the diversity of the particles, meaning that the particles cannot often move off this local optimal solution (Feng et al., 2021). To resolve this issue, an improved PSO (IPSO) method is proposed by (1) employing a dynamic inertia weight adjustment approach based on the Versoria function and (2) incorporating an adaptive mutation operation in the iterative calculation process. The Schwefel 1.2 function was adopted in this study to validate the performance of the improved PSO algorithm. The details for these two improvements are given in the following.

Step 1. Inertia weight adjustment based on the Versoria function

The inertia weight is a constant in the conventional PSO algorithm. However, previous studies have shown that a dynamic inertia weight coefficient can achieve superior optimization performances because a higher inertia weight at the iteration commencement prevents the global search from slipping into a local optimization too early. However, before the end of the iterations, a reduced inertia weight should be employed to guarantee that the algorithm converges faster and does not bounce out of the optimal solution. For example, Shi et al. (Shi and Eberhart, 1998) adopted a dynamic adjustment linear decreasing strategy approach. Although the inertia weight under this strategy is no longer constant, its change rate is still a constant, which implies that this algorithm has the same capacity as global optimization and local optimization but may not be able to achieve the global optimal solution. Chen et al. (2006) proposed a quadratic function–based approach for diminishing the inertia weight. However, although the inertia weight change rate based on this strategy changes continuously as the iteration times increase, the inertia weight change rate in the early stage is faster and then becomes slower in the later stage, which suggests that the algorithm may not only have low efficiency in local optimization but may also prematurely fall into a local optimal.

Therefore, in consideration of the above defects, this study adopts a dynamic adjustment approach based on the Versoria function for the inertia weight, which is defined as follows:

\[
f(x) = \frac{a^3}{a^2 + x^2}
\] (9)

where parameter \(a\) is the radius of the adjacent circle of the function; the larger the \(a\), the greater its descent rate between 0 and 1, and when parameter \(a\) is set to 1, it is the standard Versoria function. Considering the global optimization and convergence speed, the parameter \(a\) value in this paper is set at 2. Consequently, the following adaptive inertia weight adjustment formula is obtained:

\[
w(i) = w_{\text{min}} + (w_{\text{max}} - w_{\text{min}}) \left( \frac{12}{(\delta_i/\delta_{\text{max}})^2 + 3} - 3 \right)
\] (10)

where \(w(i)\) is the inertia weight coefficient at the \(i\)th iteration, \(w_{\text{max}}\) and \(w_{\text{min}}\) are their respective maximum and minimum values, which are respectively set at 0.9 and 0.4 in this study, \(\delta_i\) is the current number of iterations, and \(\delta_{\text{max}}\) represents the total number of iteration. The Versoria function curve and the decreasing inertia weight curve are shown in Fig. 6.

It is obvious from Fig. 6 that the preliminary function change rate is slow, which is conducive to global optimization in the solution space, while the later change rate is

![Fig. 6 Versoria function and its decreasing curve](image)
rapid, which promotes local search efficiency; therefore, the algorithm has both excellent global optimization and high convergence efficiency.

Step 2. Adaptive mutation operation

To tackle the problem associated with the progressive decline in the variety of the particle population in the iterative computation, Gandelli et al. (2007) proposed a GA-PSO hybrid optimization algorithm based on various hybrid strategies (static, dynamic, alternating, and adaptive mutation), which were verified using multimodal benchmark problems. Therefore, based on these results, this paper adds an adaptive mutation operation to the PSO iteration process. The adaptive mutation probability is calculated as follows:

$$\text{rand} > \frac{1}{2} \left( 1 + \frac{\delta_i}{\delta_{\text{max}}} \right)$$  \hspace{1cm} (11)

where in (11), \(\text{rand}\) is a random number between 0 and 1, \(\delta_i\) is the current number of iterations, and \(\delta_{\text{max}}\) is the total number of iterations. Initially, the particles escape from the group optimization direction with a certain probability, and as the evolutionary algebra increases, the particle mutation likelihood progressively decreases. This operation effectively boosts the population diversity and mitigates the risk of particles falling into a local optimum solution.

Performance evaluation of IPSO algorithm

The Schwefel 1.2 function (Yaghoobi and Esmaeili, 2017) is utilized as the benchmark function to assess the efficiency of IPSO in this work. As its independent variable has an epistatic effect, its gradient direction does not change along the axis, resulting in high optimization complexity. The independent variable for this function ranges from \(-100\) to 100. When \(x^*\) takes 0, the function has a global minimum \(f(x^*) = 0\):

$$f(x) = \sum_{i=1}^{D} \left( \sum_{j=1}^{i} x_j \right)^2$$  \hspace{1cm} (12)

To evaluate the optimization performance of the PSO and IPSO, several initialization particle swarm parameters were adjusted to be the same and then iterated 1000 times, with their respective fitness function values compared at the end of the iteration.

As shown in Fig. 7, the PSO reaches local optimization before the 700th iteration, the PSO and IPSO have the same optimization efficiency around the 920th iteration, and after that, the IPSO is still approaching the global minimum. Although the IPSO has yet to converge after 1000 iterations, the global optimization capability of the IPSO is still superior, demonstrating the effectiveness of the improvements over the standard PSO.

LSTM based on IPSO optimization (IPSO-LSTM)

As the LSTM layers contain several hyperparameters, such as the learning rate, the number of hidden units, and the epoch times, any slight changes to certain parameters may result in substantial differences in the fitting outcomes. Because there is a huge number of combinations of hyperparameter variables, it is challenging to identify the best answers for an MSW generation forecasting model based on deep learning methods (Kim and Cho, 2021). Consequently, this study optimizes the forecasting model’s performance using the IPSO algorithm and selects the mean square error reciprocal for the training set as the fitness function, which automatically searches the hyperparameter values for the LSTM neural network using iterative calculation and information sharing between the populations.
The IPSO-LSTM model parameter settings are divided into two parts: the IPSO parameter initialization and the LSTM parameter search interval. The specific parameter settings are shown in Table 2.

In Table 2, for the IPSO parameters, PN denotes the population size for the proposed algorithm, MAXITER is the number of iterations, \( c_1 \) and \( c_2 \) are respectively the individual learning factor and group learning factor, \( r_1 \) and \( r_2 \) are random numbers ranging from 0 to 1, and \( w_{\min} \) and \( w_{\max} \) are the minimum and maximum inertia weight values. For the LSTM parameters, LR is the learning rate search range, EPOCH is the search range for the neural network iterations, and NFHL and NSHL respectively indicate the search ranges for the number of nodes in the first and second hidden layers.

## Improved GM (1,1) model

The Grey system theory, which was proposed by Deng in 1981 (Wang et al., 2010), can utilize a small quantity of data and imperfect information to anticipate and regulate a system’s future trends and situations. As the GM (1,1) requires a small amount of data, is easy to operate, and obtains good forecast performances, it has become the most commonly used grey forecasting model and has been widely used in various fields of study because of its excellent prediction performances.

The fundamental idea of the GM (1,1) model is to test the stationarity of the original series, accumulate and generate new series to weaken the randomness, and establish corresponding differential equation models to reveal the characteristic laws for the new series so that an orderly sequence analysis can be performed and the predictions can be fulfilled. However, the predictive precision of the original GM (1,1) model needs to be enhanced because errors can originate from the background value, the initial value selection, and the data stationarity (Liu et al., 2014). Therefore, to minimize the errors associated with the background value optimization strategy, this study conducted a translation transformation on the original sequence to enhance its stability and diminish the evident modeling process system errors.

### Translation transformation

Translation transformations test the known sequences to ensure conformity in the modeling method feasibility. Set the original sequence as \( x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)) \), and then calculate the stepwise ratio:

\[
\lambda(k) = \frac{x^{(0)}(k-1) - x^{(0)}(k)}{x^{(0)}(k)}, k = 2, 3, \ldots, n
\]  

This sequence can then be used for grey prediction if all stepwise ratios \( \lambda(k) \) fall within the allowable coverage \( \Theta = \left( e^{-\frac{2}{k}}, e^{\frac{2}{k}} \right) \). Otherwise, it needs to be further transformed to improve the forecasting accuracy. In this experiment, an appropriate constant \( c \) was added to each original sequence number to ensure it passed the stepwise ratio test:

\[
\begin{align*}
\lambda_{c}(k) &= \frac{y^{(0)}(k) - c}{y^{(0)}(k)}, k = 2, 3, \ldots, n \\
\lambda_{c}(k) &= \frac{y^{(0)}(k) - c}{y^{(0)}(k)} \in \Theta, k = 2, 3, \ldots, n
\end{align*}
\]  

### Background value optimization based on error minimization

The GM (1,1) time response formula demonstrates that the model’s prediction accuracy is highly reliant on the development coefficient (\( a \)) and gray control variable (\( b \)) calculations, which are contingent on precise background value calculations. However, one of the prime causes of conventional GM (1,1) inaccuracy is approximate background value calculations (Xu et al., 2015). The background value optimization method based on error minimization proposed by Xu et al. (2015) provides a direction for decreasing these system errors. Therefore, based on this idea, the integral mean value theorem was applied to the real background value, and the least square method was used to estimate the grey differential equation parameters and ensure that the background value had both unbiased and minimum errors.

Based on the above analysis, a linear combination background value construction form is adopted; therefore, the GM (1,1) grey differential equation model was amended as follows:

\[
x^{(0)}(k) + \alpha [a x^{(1)}(k - 1) + (1-a) x^{(1)}(k)] = b
\]  

where \( x^{(0)} \) is the original sequence, \( x^{(1)} \) is the sequence generated by its accumulation, \( a \) is the development coefficient, and \( b \) is the gray control variable. Therefore, \( r = (a^*, b^*)^T \) are the parameters to be solved in the model, which can be obtained using the least-squares method. The estimated \( a, b, \) and \( \alpha \) parameter values using the least-squares method are:

### Table 2 IPSO-LSTM parameter settings

| IPSO | LSTM |
|------|------|
| PN   | MAXITER | [c_1, c_2] | [r_1, r_2] | [w_{\max}, w_{\min}] | LR | EPOCH | NFHL | NSHL |
| 10   | 20     | [1.5, 2]   | [0.8, 0.3] | [0.9, 0.4] | [0.005, 0.01] | [100,1000] | [1.50] | [1.50] |
Substituting the above parameters into the whitening differential equation yields the discrete time response formula:

\[
\hat{x}^{(1)}(k) = \frac{b}{a} + \left(\frac{\hat{x}(0)}{a} - \frac{b}{a}\right)e^{-a(k-1)}
\]  

(17)

After the progressive reduction, the forecast sequence is obtained as follows:

\[
\begin{align*}
\alpha &= \frac{1}{\ln(\hat{a}^* + 1)} - \frac{1}{\hat{a}^*}, \\
\alpha &= \ln(\hat{a}^* + 1), \\
b &= \frac{b^*\ln(\hat{a}^* + 1)}{\hat{a}^*}.
\end{align*}
\]  

(16)

**Forecasting model construction**

To investigate the potential evolutionary impact of the selected factors under different policy orientations and social development backgrounds on future MSW generation trends, a scenario analysis method based on a hybrid forecasting model was adopted to predict the MSW generation in Shanghai from 2025 to 2035. As shown in Fig. 8, the hybrid forecasting model proposed in this paper was divided...
into three phases: forecasting the input variables, constructing the causal relationship forecasting model, and scenario forecasting.

The first phase considered the future socio-economic development possibilities, for which the growth trends for each influencing factor were divided into three scenarios and the economic and social factors predicted using the improved GM (1,1). For the medium scenario, the index growth was based on normal economic and social development. As the CUPS established a defined plan for the permanent resident population and the number of overseas tourists, the demographic factors are expected to rise linearly and steadily to a peak. The development trends for each influencing factor under the high and low scenarios were calculated based on the medium scenario to demonstrate the growth trends under various circumstances, such as diversified policy orientations, socio-economic development, and demographic changes.

For the second phase, a causal relationship forecasting model was developed based on the IPSO-LSTM, which was then employed to fit the highly nonlinear relationships between the nine influencing factors and the MSW generation. As the LSTM neural network has many hyperparameters to be adjusted, manual parameter adjustment would not obtain an optimal parameter set; therefore, the IPSO algorithm was adopted to search for the hyperparameter combinations that minimized the sum of mean squared error (MSE) and L2 regularization term of the training set. If the mean absolute percentage error (MAPE) of the test set fulfills the assessment requirements given in the literature (Table 3), it indicates that the model performance is excellent and can be utilized for subsequent multifactor prediction.

Finally, in the third phase, with the predicted values for the influencing factors determined as the input of IPSO-LSTM model, the MSW generation predictions under the three scenarios were determined, and the three-phase causal relationship forecasting method based on the IPSO-LSTM was constructed. The prediction model organically integrates the time series prediction method with the causal relationship forecasting method to alleviate the single prediction method limitations, considers the possible diverse MSW future generation growth trends, and pays attention to the factors and coordination relationship analyses that are crucial to the system’s development.

### Table 3 The criteria of MAPE

| MAPE(%) | Forecasting power |
|---------|-------------------|
| < 10    | Excellent         |
| 10–20   | Good              |
| 20–50   | Reasonable        |
| > 50    | Incorrect         |

### Proposed model performance test

**IPSO-LSTM model performance evaluation**

Based on the real-world Shanghai data from 1980 to 2019, the optimized LSTM neural network hyperparameters, LR, EPOCH, NFHL, and NSHL, are shown in Table 4. To test the performance of this hyperparameter combination, the following five evaluation metrics are proposed.

**Evaluation metrics**

To evaluate the IPSO-LSTM forecasting accuracy, mean absolute percentage error (MAPE), root mean square error (RMSE), mean absolute error (MAE), the coefficient of determination ($R^2$), and the Theil inequality coefficient (TIC) were chosen as the assessment indicators, the relevant formulas for which were as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \quad (20)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \quad (21)$$

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (22)$$

$$TIC = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} \hat{y}_i^2} + \sqrt{\frac{1}{n} \sum_{i=1}^{n} y_i^2}} \quad (23)$$

where $y_i$ was the actual value, $\hat{y}_i$ was the forecast, and $\bar{y}_i$ was the average value of the dataset. MAPE, $R^2$, and TIC generally vary from 0 to 1, with the closer $R^2$ is to 1, the greater the trend becomes between the forecast value and the actual value, and the lower the TIC value, the smaller the gap between the forecast and the actual values, indicating better forecasting precision. The RMSE and MAE range

### Table 4 Optimal LSTM hyperparameters

| LR     | EPOCH | NFHL | NSHL |
|--------|-------|------|------|
| 0.00092| 1061  | 2    | 19   |
from 0 to positive infinity, with the smaller the RMSE and MAE values, the better. For MAPE, Hsu and Wang (2007) proposed the corresponding assessment approach:

**IPSO-LSTM prediction effectiveness**

The data from the first 32 years were taken as the training set and the subsequent 8 years were taken as the test set. These were then respectively fed into the IPSO-LSTM model to determine the prediction results, which are shown in Fig. 9 and Table 5.

The iterative training set and the test set operations converged the errors and determined the optimal solution. The evaluation metrics also found that the model had an excellent fitting ability and generalization capacity, particularly at the mutation points, which meant that based on the future influencing factors, it could be applied to predict future MSW generation.

**Comparative verification**

Back propagation neural network (BPNN) and support vector regression (SVR) are two widely used models in the MSW forecasting area; thus, to further verify the performance of the proposed hybrid model, the following two comparisons are performed: (1) the first comparison is conducted to show the good performance of LSTM model compared to BPNN and SVR methods, and (2) the second comparison is performed to investigate the influence of IPSO on the LSTM neural network.

Table 6 shows the error outcomes, and Fig. 10 shows the error radar plots for the five distinct methods, in which the greater the comprehensive forecast effectiveness, the smaller the region in the radar graphic. As can be seen, the LSTM prediction accuracy was higher than the BPNN and SVR, which verified that compared with traditional neural networks and the other nonlinear regression methods, the LSTM has excellent generalization ability and can fit highly nonlinear systems. The IPSO-LSTM precision was found to be better than the LSTM precision, which suggested that the IPSO optimized hyper-parameters allow the LSTM to thoroughly exploit its superior prediction ability.

**Improved GM (1,1) method performance**

**Evaluation metrics**

Posterior error detection method criteria including small error probability \(p\) and the post-error ratio \(C\); were employed to assess the GM (1,1) accuracy. Under the assumption that \(S_1\) and \(S_2\) were the respective standard deviations for the original sequence \(X^{(0)}\) and the error sequence \(\varepsilon^{(0)}\), \(S_1\) and \(S_2\) were calculated as follows:

\[
S_1 = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( X^{(0)}(i) - \frac{1}{n} \sum_{i=1}^{n} X^{(0)}(i) \right)^2} \quad (24)
\]

\[
S_2 = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \varepsilon(i) - \frac{1}{n} \sum_{i=1}^{n} \varepsilon(i) \right)^2} \quad (25)
\]

The small error probability \(p\) and post-error ratio \(C\) were then calculated as follows respectively:
where the \( p \)-value was the error fraction within the acceptable range, and the \( C \) value was the change rate for the difference between the predicted and the actual values (Yousuf et al., 2021). Therefore, the greater the \( p \)-value and the lower the \( C \) value, the better the forecasting performance. The acceptable \( p \) and \( C \) levels are shown in Table 7.

### Comparative verification

To verify its effectiveness, the IGM(1,1) was then compared with the original GM (1,1) based on the above evaluation metrics, with the results shown in Table 8, which shows that the IGM(1,1) has good prediction error acceptability and only a small gap between the actual and forecast values.

### Shanghai MSW generation scenario analyses

Scenario analysis is a qualitative prediction approach that assesses the prediction object based on the assumption that a specific trend is expected to continue in the future (Zhang et al., 2021). The scenario analysis approach which considers as many alternatives as possible is more feasible than single-point forecasting methods and provides decision-makers with comprehensive decision-making schemes. Therefore, scenario analysis has been widely used for energy and environmental modeling.

#### Scenario divisions and factor forecasting

**Scenario setting**

As this study intended to simulate the impact of multiple influencing factors on MSW generation under different growth tendencies, the scenario analysis needed to adequately address the probable social, economic, and demographic developments. Given this context, the Shanghai MSW generation from 2025 to 2035 was predicted for three separate scenarios. As the CUPS explicitly stated that Shanghai’s permanent residents would be controlled at 25 million and the number of overseas tourists was predicted to be 14 million by 2035, these two indicators were independently analyzed. The improved GM (1,1) was then applied to evaluate the additional factors not expressly noted in this plan. The scenario definitions, scenario descriptions, and the definition basis were as follows:

**I. Medium scenario** The medium scenario is constructed based on the national 14th Five-Year Plan and the CUPS. In this scenario, it is expected that Shanghai would strictly control its urban scale and control the resident population growth to 25 million. It would also exploit its urban tourist resources to nurture and develop high-quality tourism service functions to become a world-renowned tourism destination city, with the number of overseas tourists increasing to 14 million by 2035. As the core city of the world-class Yangtze River Delta urban agglomeration, Shanghai would take full advantage of its development potential and aggressively explore new economic models. Therefore, the GDP, TI, CPI, RPI, UPCCE, and UPCDI economic factors would generate new growth vitality. Socially, Shanghai would seek to improve people’s livelihoods by expanding their livelihood benefits and continuing to steadily increase its investment in urban infrastructure construction. Under this medium scenario, the improved GM (1,1) was used to anticipate the

---

**Table 6** Comparison of the prediction performances of the different models

| Methods     | MAPE (%) | RMSE    | MAE     | \( R^2 \) | TIC    |
|-------------|----------|---------|---------|----------|--------|
| IPSO-LSTM   | 6.28     | 60.2500 | 52.3602 | 0.7178   | 0.0351 |
| ARIMA       | 6.81     | 83.3343 | 63.0699 | 0.4601   | 0.0502 |
| LSTM        | 7.23     | 82.0535 | 65.6662 | 0.4766   | 0.0492 |
| SVR         | 7.50     | 101.2884| 89.4319 | 0.0030   | 0.0685 |
| BPNN        | 8.54     | 90.9766 | 76.1667 | 0.3566   | 0.0544 |

**Table 7** GM(1,1) accuracy testing

| Accuracy level | Small error probability \((p)\) | Post-error ratio \((C)\) |
|----------------|---------------------------------|------------------------|
| I              | \(0.95 \leq p<1\)                | \(C \leq 0.35\)         |
| II             | \(0.80 \leq p<0.95\)             | \(0.35 < C \leq 0.50\)  |
| III            | \(0.70 \leq p<0.80\)             | \(0.50 < C \leq 0.65\)  |
| IV             | \(0.60 \leq p<0.70\)             | \(0.65 < C \leq 0.80\)  |

\[
p = P\{\left| \varepsilon(k) - \bar{\varepsilon} \right| < 0.6745S_1 \} \quad (26)
\]

\[
C = \frac{S_2}{S_1} \quad (27)
\]

where the \( p \)-value was the error fraction within the acceptable range, and the \( C \) value was the change rate for the difference between the predicted and the actual values (Yousuf et al., 2021). Therefore, the greater the \( p \)-value and the lower the \( C \) value, the better the forecasting performance. The acceptable \( p \) and \( C \) levels are shown in Table 7.

**Fig. 10** Radar map of the errors
economic and social development factors because these were not clearly described in the relevant planning documents.

II. Low scenario  Demographic factors have been identified as the primary contributors to MSW generation (Ayeleru et al., 2021). However, in the next 25 years, China is expected to become an aging society at a faster pace than Japan and some western developed countries (Chen et al., 2019), which means there would be a decrease in per capita consumption and the workforce proportion and a reduction in per capita MSW generation. It is also expected that the repercussions from the continuing COVID-19 pandemic would result in a sustained global economic downturn, which means that Shanghai’s economic development would inevitably suffer because of its status as the free trade port with one of the world’s largest import and export quantities. The post-COVID-19 era consumption psychology would also lead to more conservative consumption behavior (Ren and Guo, 2020). In this context, the development trends for the relevant influencing factors under this low scenario would have deteriorating trends and be under great pressure. The low scenario, therefore, sees the resident population grow to only 20 million, the number of overseas tourists grows to only around 10 million, and the other influential factors decline by 10% compared with the medium scenario.

III. High scenario To promote the long-term balanced development of the population and optimize its fertility policy, China would fully liberalize its population policy by implementing a three-child fertility policy and corresponding supporting measures, which would alleviate the population aging, increase the workforce proportion and consumption capacity, and indirectly lead to a rise in the per capita MSW generation. China’s 14th Five-Year Plan and its long-term objectives for 2035 explicitly stated that it would accelerate the establishment of a “dual circulation” development pattern in which the domestic economic cycle plays a leading role while the international economic cycle remains its extension and supplement. As the metropolis at the forefront of economic aggregate and comprehensive urban functions, to advance people’s livelihoods, Shanghai would deepen reforms, optimize its economic structure, execute a high-level opening to the outside world, strengthen its public service system, and enhance its urban management system. As the associated infrastructure construction would be enhanced under this planning pattern, the overall fixed assets investment would increase. Therefore, under the high scenario, there would be a significant increase in the relevant influencing factors, with the resident population growing to 30 million, the number of foreign tourists growing to 20 million, and the other influencing factors being 10% greater than in the medium scenario.

Factor forecasting under the different scenarios

Based on the above three scenario development trends, the anticipated influencing factor values from 2025 to 2035 were respectively determined, the results for which are shown in Fig. 11:

Shanghai MSW forecasts from 2025 to 2035

The annual MSW generation from 2025 to 2035 was determined by inputting the forecast results for the annual influencing factors under the different scenarios into the IPSOLSTM model, the results for which are shown in Fig. 12.

The projection outcomes showed that the MSW generation under all three scenarios would rise every year; however, the growth rate would be somewhat slower than the historical trend in the recent 5 years, which could be because of the strict control of Shanghai’s urban scale and resident population. It was also observed from Fig. 13 that the MSW generation gap between the high and medium scenarios was decreasing each year, which indicated that although the primary influencing factor, population growth, would be limited by the urban scale planning, the continuous economic and social development would result in higher personal consumption levels and social service capacity, which would indirectly lead to an increase in MSW generation.

Because the analysis indicated that the MSW generation under the medium scenario would move closer to the high scenario over time, to achieve its long-term goal of harmonious coexistence between humans and nature and the building of an ecologically sound, socialist modern international metropolis with world influence by 2035, to deal with a possible environmental waste crisis in the future, Shanghai would need to take the following whole life cycle

Table 8  Comparative verification of improved GM (1,1)

|          | GDP  | TI   | CPI  | RPI  | UPCCE | UPCDI | TIFA |
|----------|------|------|------|------|-------|-------|------|
| Original GM (1,1) |      |      |      |      |       |       |      |
| p        | 0.525| 0.900| 1.000| 0.700| 0.900 | 0.850 | 0.875|
| C        | 1.027| 0.473| 0.339| 0.539| 0.438 | 0.601 | 0.494|
| Improved GM (1,1) |      |      |      |      |       |       |      |
| p        | 1.000| 1.000| 1.000| 0.875| 0.950 | 0.975 | 1.000|
| C        | 0.312| 0.297| 0.244| 0.486| 0.325 | 0.290 | 0.182|
waste management measures including waste classification, infrastructure investment, digital management, and resource utilization. The detailed measures are given as follows.

1. Strengthen citizen awareness of waste categorization and increase investment in municipal infrastructure construction.

A crucial requirement for subsequent resource utilization is effective waste categorization. The government and social institutions should adopt multiple information channels, such as television, internet, school education, and advertising, to raise public waste classification, recycling, and green habit awareness. Because the prediction results indicated that the MSW generation under the medium scenario would approach that of the high scenario, the planning and construction of municipal infrastructures, such as waste transfer stations, recycling bins, and landfill sites, need to be increased to alleviate the pressure on the current waste recycling and treatment systems and improve overall MSW disposal efficiency.

2. Realize efficient MSW management by integrating digitization and informatization.

Fig. 11 Prediction results of influencing factors under three scenarios

Fig. 12 Scenario prediction results

| Year | High Scenario | Medium Scenario | Low Scenario |
|------|---------------|----------------|-------------|
| 2025 | 10.0          | 9.7            | 9.1         |
| 2026 | 10.1          | 9.8            | 9.2         |
| 2027 | 10.0          | 10.0           | 9.3         |
| 2028 | 10.4          | 10.2           | 9.5         |
| 2029 | 10.6          | 10.3           | 9.6         |
| 2030 | 10.7          | 10.5           | 9.7         |
| 2031 | 10.8          | 10.7           | 9.8         |
| 2032 | 11.0          | 10.8           | 9.9         |
| 2033 | 11.1          | 11.0           | 10.0        |
| 2034 | 11.3          | 11.1           | 10.1        |
| 2035 | 11.4          | 11.3           | 10.2        |
Digitization and informatization can coordinate and connect materials and information flow using automated platforms and the Internet of Things (IoT) to categorize, sort, or recycle consumer and business waste. Therefore, to ensure more sustainable resource consumption, developing a large city-wide data hub could increase industrial productivity, minimize overproduction, decrease waste at the generation source, and also allow for the more effective recovery and recycling of MSW secondary materials.

3. Adopt emerging technologies for waste-to-energy and waste-to-material.

Shanghai’s whole life MSW treatment cycle should adopt sophisticated and mature technologies, such as material recovery with composting, for resource utilization and final stage harmless treatment to decrease its environmental impact. Compared with landfill disposal, which has been widely accepted and applied, thermal treatments such as incineration, gasification, and pyrolysis and biological treatments such as aerobic composting and anaerobic digestion not only have less secondary pollution but can also convert MSW into fertilizer, biogas, biofuel, electricity, and other resources.

Conclusions

To achieve its long-term goal of becoming an outstanding global city and a socialist modern international metropolis with global influence by 2035, Shanghai must coordinate its municipal infrastructures and modernize its comprehensive solid waste treatment capacity; however, these goals are reliant on accurate MSW generation forecasts. To fully consider Shanghai’s urban, economic, and social development plans, three possible future scenarios were developed, to which the proposed IPSO-LSTM model was applied to predict the future MSW generation and provide suggestions for the development of an effective future MSW treatment system in Shanghai.

The model analysis revealed that the IPSO significantly enhanced the fitting performance of the LSTM neural network, especially at the mutation points, and compared with the LSTM, BPNN, and SVR reference models, the IPSO-LSTM was found to have better prediction performances. It was also found that the proposed IGM(1,1) employed in this research outperformed the original GM(1,1) in terms of prediction performances and medium- and long-term prediction efficiency. Finally, compared with single time series forecasting or causal relationship forecasting methods, as the IPSO-LSTM and IGM(1,1) scenario prediction method proposed in this paper accounted for the possible diverse future MSW generation growth trends by analyzing the factor coordination relationships crucial to the MSW system development, it overcame the shortcomings associated with single methods.

The experimental results showed that MSW generation from 2025 to 2035 would continue to increase under all three scenarios, with the maximum generation range being 10.24034 million tons to 11.40179 million tons in 2035. Currently, there are 11 waste treatment facilities of different specifications in Shanghai, which together have a daily treatment capacity of 1.39225 million tons and an annual treatment capacity of around 508.2 million tons. Therefore, by 2035, the existing waste treatment system would need to operate at 2.22 to 2.24 times higher than the designed scale, which would accelerate equipment depreciation and pose a significant threat to the safe waste treatment system operations. Therefore, to decrease the MSW generation, greater investment is needed in the waste treatment system, the public’s waste categorization awareness must be raised.
and informatization and digitalization must be incorporated into the waste management systems.

This study provided a novel MSW generation forecasting method for authorities and academics to reveal the long-term MSW generation implications, which has rarely been examined in previous literature and can hopefully inspire further research in the field. However, the limitations of this work were the possible dataset heterogeneity and the lack of data from lower administrative units. Future research could conduct surveys to collect additional MSW generation data from lower residential regions to complement or validate this algorithm and comparison MSW management strategy studies between Shanghai and other international metropolises.

Acknowledgements The authors appreciate the computing resources provided by School of Economics and Management at the China University of Geosciences.

Author contribution Deyun Wang: conceptualization, methodology, software, validation, formal analysis, supervision, writing-original draft, writing-review and editing. Ying-an Yuan: methodology, data collection and curation, software, validation, formal analysis, visualization, writing-original draft, writing-review and editing. Hongyuan Luo: writing-review and editing. Haixiang Guo: writing-review and editing.

Funding This work was supported by Hubei Provincial Natural Science Foundation of China (Grant No. 2020CFF822), National Social Science Foundation of China (Grant No. 21AZD074), National Natural Science Foundation of China (Grant Nos.71874165, 72074198), National College Students’ innovation and entrepreneurship training program (Grant No. 20210491031).

Data availability The datasets used and analyzed in this study are available from Shanghai statistical yearbooks (http://tjj.sh.gov.cn) and the National statistical yearbooks (http://www.stats.gov.cn).

Declarations

Ethical approval Not applicable.

Consent to participate All authors were participated in this work.

Consent for publication All authors allow the publication of this paper.

Competing interests The authors declare no competing interests.

References

Abdallah M, Abu Talib M, Feroz S, Nasir Q, Abdalla H, Mahfood B (2020) Artificial intelligence applications in solid waste management: a systematic research review. Waste Manage 109:231–246. https://doi.org/10.1016/j.wasman.2020.04.057

Araiza-Aguilar JA, Rojas-Valencia MN, Aguilar-Vera RA (2020) Forecast generation model of municipal solid waste using multiple linear regression. Glob J Environ Sci Manag 6(1):1–14. https://doi.org/10.22034/GJESM.2020.01.01

Ayeleru OO, Fajimi LI, Oboirien BO, Olubambi PA (2021) Forecasting municipal solid waste quantity using artificial neural network and supported vector machine techniques: a case study of Johannesburg, South Africa. J Clean Prod 289:125671. https://doi.org/10.1016/j.jclepro.2020.125671

Chang NB, Lin YT (1997) An analysis of recycling impacts on solid waste generation by time series intervention modeling. Resour Conserv Recycl 19(3):165–186. https://doi.org/10.1016/S0921-3449(96)01187-1

Chao YL (2008) Time series analysis of the effects of refuse collection on recycling: Taiwan’s “Keep Trash Off the Ground” measure. Waste Manage 28(5):859–869. https://doi.org/10.1016/j.wasman.2007.02.012

Chen EI, Hewel J, Felding-Habermann B, John RY (2006) Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT)*. Mol Cell Proteom 5(1):53–56. https://doi.org/10.1074/mcp.T500013-MCP200

Chen HW, Chang NB (2000) Prediction analysis of solid waste generation based on grey fuzzy dynamic modeling. Resour Conserv Recycl 29(1):1–18. https://doi.org/10.1016/S0921-3449(99)00052-X

Chen R, Xu P, Song PP, Wang MF, He JJ (2019) China has faster pace than Japan in population aging in next 25 years. Biosci Trends 13(4):287–291. https://doi.org/10.5582/bst.2019.01213

Deng Y, Fu Y, Zhang H, Liu X, Liu Z (2022) Protein post-translational modification site prediction using deep learning. Procedia Comput Sci 198:480–485. https://doi.org/10.1016/j.procs.2021.12.273

Feng H, Ma W, Yin CB, Cao DH (2021) Trajectory control of electro-hydraulic position servo system using improved PSO-PID controller. Autom Constr 127:103722. https://doi.org/10.1016/j.autcon.2021.103722

Gandelli A, Grimaccia F, Mussetta M, Pirinoil P (2007) Development and validation of different hybridization strategies between GA and PSO. 2007 IEEE Congress on Evolutionary Computation VOLS 1–10, PROCEEDINGS, 2782–+. https://doi.org/10.1109/CEC.2007.4424823

Ghinea C, Drăgoi EN, Comanăţiu ED, Gavrilescu M, Câmpian T, Curtanu S, Gavrilescu M (2016) Forecasting municipal solid waste generation using prognostic tools and regression analysis. J Environ Manage 182:80–93. https://doi.org/10.1016/j.jenvman.2016.07.026

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hsu LC, Wang CH (2007) Forecasting the output of integrated circuit industry using a grey model improved by the Bayesian analysis. Technol Forecast Soc Chang 74(6):843–853. https://doi.org/10.1016/j.techecon.2006.02.005

Kannangara M, Dua R, Ahmadi L, Benebah F (2018) Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches. Waste Manage 74:3–15. https://doi.org/10.1016/j.wasman.2017.11.057

Kim TY, Cho SB (2021) Optimizing CNN-LSTM neural networks with PSO for anomalous query access control. Neurocomputing 456:666–677. https://doi.org/10.1016/j.neucom.2020.07.154

Kolekar KA, Hazra T, Chakraborty SN (2016) A review on prediction of municipal solid waste generation models. Procedia Environ Sci 35:238–244. https://doi.org/10.1016/j.proenv.2016.07.087

Lin K, Zhao Y, Tian L, Zhao CL, Zhang ML, Zhou T (2021) Estimation of municipal solid waste amount based on one-dimension convolutional neural network and long short-term memory with attention mechanism model: a case study of Shanghai. Sci Total Environ 791:148088. https://doi.org/10.1016/j.scitotenv.2021.148088
Shahid F, Zameer A, Muneeb M (2021) A novel genetic LSTM model.

Niu DJ, Wu F, Dai SJ, He S, Wu BR (2021) Detection of long-term health status of children based on machine learning and spatio-temporal clustering. Expert Syst Appl 57:342–351. https://doi.org/10.1016/j.eswa.2016.06.011

Nguyen XC, Nguyen TTH, La DD, Kumar C, Rene ER, Nguyen DD, Chang SW, Chung WJ, Nguyen XH, Nguyen VK (2021b) Development of machine learning based models to forecast solid waste generation in residential areas: a case study from Vietnam. Resour Conserv Recycl 167:105381. https://doi.org/10.1016/j.resconrec.2020.105381

Niu DJ, Wu F, Dai SJ, He S, Wu BR (2021) Detection of long-term effects of forecasting municipal solid waste using a long short-term memory network in a study in Vietnam. J Clean Prod 320:125187. https://doi.org/10.1016/j.jclepro.2021.125187

Ren FF, Guo RJ (2020) Public mental health in post-COVID-19 era. Psychiatr Danub 32(2):251–255. https://doi.org/10.24869/psyd.2020.251

Shahid F, Zameer A, Muneeb M (2021) A novel genetic LSTM model for wind power forecast. Energy 223:120069. https://doi.org/10.1016/j.energy.2021.120069

Shi Y and Eberhart R (1998) A modified particle swarm optimizer. 1998 IEEE International Conference on Evolutionary Computation. Proceedings. IEEE World Congress on Computational Intelligence, Cat. No.98TH8360. https://doi.org/10.1109/ICEC.1998.699146

Solano Meza JK, Orjuela Yepes D, Rodrigo-Illarri J, Cassiraga E (2019) Predictive analysis of urban waste generation for the city of Bogotá, Colombia, through the implementation of decision trees-based machine learning, support vector machines and artificial neural networks. Heliyon 5(11):e02810. https://doi.org/10.1016/j.heliyon.2019.e02810

Sun QQ, Ge ZQ (2021) Deep learning for industrial KPI prediction: when ensemble learning meets semi-supervised data. IEEE Trans Industr Inf 17(1):260–269. https://doi.org/10.1109/TII.2020.2969709

Sunayana KS, Kumar R (2021) Forecasting of municipal solid waste generation using non-linear autoregressive (NAR) neural models. Waste Manage 121:206–214. https://doi.org/10.1016/j.wasman.2020.12.011

Wang F, Yu L, Wu A (2021) Forecasting the electronic waste quantity with a decomposition-ensemble approach. Waste Manage 120:828–838. https://doi.org/10.1016/j.wasman.2020.11.006

Wang YH, Dang YG, Li YQ, Liu SF (2010) An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. Expert Syst Appl 37(8):5640–5644. https://doi.org/10.1016/j.eswa.2010.02.048

Xiao SJ, Dong HJ, Geng Y, Francisco MJ, Pan HY, Wu F (2020) An overview of the municipal solid waste management modes and innovations in Shanghai, China. Environ Sci Pollut Res 27(24):29943–29953. https://doi.org/10.1007/s11356-020-09398-5

Xu LL, Gao P, Cui SH, Liu C (2013) A hybrid procedure for MSW generation forecasting at multiple time scales in Xiamen City, China. Waste Manag 33(6):1324–1331. https://doi.org/10.1016/j.wasman.2013.02.012

Xu N, Dang YG, Ding S (2015) Optimization method of background value in GM(1,1) model based on least error. Control Decis 30(02):283–288. https://doi.org/10.13195/j.kzyjc.2013.1808

Yaghoobi T, Esmaeili E (2017) An improved artificial bee colony algorithm for global numerical optimisation. Inf Proc Lett 120(4):251–258. https://doi.org/10.1016/j.ipl.2011.06.002

Yan R, Liao J, Yang J, Sun W, Nong MY, Li FP (2021) Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Syst Appl 169:114513. https://doi.org/10.1016/j.eswa.2020.114513

Yousuf MU, Al-Bahadly I, Avci E (2021) A modified GM(1,1) model to accurately predict wind speed. Sustain Energy Technol Assess 43:100905. https://doi.org/10.1016/j.seta.2020.100905

Yu L, Wang S, Lai KK (2008) Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ 30(5):2623–2635. https://doi.org/10.1016/j.eneco.2008.05.003

Yuan XF, Li L, and Wang YL (2020) Nonlinear dynamic soft-sensor modeling with supervised long short-term memory network. IEEE Trans Industr Inf 16(5):3168-3176. https://doi.org/10.1109/TII.2019.2902129

Zhang F, Chan APC, Darko A, Li DZ (2021) Predicting the elderly’s health status using machine learning based on spatio-temporal data. Int J Environ Res Public Health 18(18):10095-10095. https://doi.org/10.3390/ijerph181810095

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.