International Journal of Mathematical, Engineering and Management Sciences
Vol. 6, No. 5, 1276-1296, 2021
https://doi.org/10.33889/IJMEMS.2021.6.5.077

Application of Modified Grey Forecasting Model to Predict the Municipal Solid Waste Generation using MLP and MLE

Mohd Anjum
Department of Computer Engineering,
Aligarh Muslim University, Aligarh, Uttar Pradesh, India.
Corresponding author: mohdanjum@zhcet.ac.in

Sana Shahab
Department of Business & Administration,
Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia.
E-mail: sshahab@pnu.edu.sa

Mohammad Sarosh Umar
Department of Computer Engineering,
Aligarh Muslim University, Aligarh, Uttar Pradesh, India.
E-mail: saroshumar@zhcet.ac.in

(Received on April 4, 2021; Accepted on August 5, 2021)

Abstract
Grey forecasting theory is an approach to build a prediction model with limited data to produce better forecasting results. This forecasting theory has an elementary model, represented as the GM(1,1) model , characterized by the first-order differential equation of one variable. It has the potential for accurate and reliable forecasting without any statistical assumption. The research proposes a methodology to derive the modified GM(1,1) model with improved forecasting precision. The residual series is forecasted by the GM(1,1) model to modify the actual forecasted values. The study primarily addresses two fundamental issues: sign prediction of forecasted residual and the procedure for formulating the grey model. Accurate sign prediction is very complex, especially when the model lacks in data. The signs of forecasted residuals are determined using a multilayer perceptron to overcome this drawback. Generally, the elementary model is formulated conventionally, containing the parameters that cannot be calculated straightforward. Therefore, maximum likelihood estimation is incorporated in the modified model to resolve this drawback. Three statistical indicators, relative residual, posterior variance test, and absolute degree of grey indices, are evaluated to determine the model fitness and validation. Finally, an empirical study is performed using actual municipal solid waste generation data in Saudi Arabia, and forecasting accuracies are compared with the linear regression and original GM(1,1). The MAPEs of all models are rigorously examined and compared, and then it is obtained that the forecasting precision of GM(1,1) model , modified GM(1,1) model, and linear regression is 15.97%, 8.90%, and 27.90%, respectively. The experimental outcomes substantiate that the modified grey model is a more suitable forecasting approach than the other compared models.

Keywords- GM(1,1) model, Linear regression, Multilayer perceptron, Maximum likelihood estimation, Forecasting.

1. Introduction

Professor Deng Julong proposed the idea of the grey system in the year 1982 to model problems with limited information (Balochian & Baloochian, 2020). This theory has numerous applications in distinct natural and social science areas, and time series prediction is a primary application. The Grey theory of forecasting is an epoch-making embranchment of grey system theory (Liu & Xie, 2019). The GM(1,1) model is an elementary grey forecasting model which is represented by the first order differential equation with single variable. Therefore, it is also called the first-order grey model. It is derived using limited data points (4 or more), but the results have adequate accuracy. So, it has been implemented in many specialized fields of economy, manufacturing, agriculture, and power with high precision results (Hu, 2020; Hu et al., 2020; Li & Zhang, 2021; Zeng et al., 2020). Municipal solid waste management (MSWM) has emerged as a critical issue from the local to the global level. It directly affects the environment, water, soil, and most critically human health. Generally, MSWM operations are administered by local government bodies, so it is the most challenging issue for civic bodies. An effective and efficient administration is possible through proper planning and resources needed for different operations involved in MSWM. The need for resources only depends on the amount of municipal solid waste (MSW) generated. The MSW generation data is started to record for a few years, so MSWM bodies lack data. Therefore, the grey forecasting model is a potential approach to predict future MSW generation. However, different variants of the grey model are utilized in diverse disciplines and exhibited potentially satisfactory results. Still, numerous studies divulge that there exists a likelihood to enhance the model performance further.

The research derives a modified grey model by incorporating residuals in original forecasted values. The sign of residual is predicted through multilayer perceptron (MLP) to add or subtract residual into corresponding forecasted value. The model parameters can be determined using a statistical method such as maximum likelihood estimation (MLE) and an optimization approaches that minimizes the error function. Nowadays, metaheuristic algorithms are most prevalent to optimize the nonlinear error function. These algorithms comprise many optimization algorithms, namely particle swarm optimization, grey wolf optimization, flower pollination, gravitational search, Cuckoo search, differential evolution, evolution strategy etc. (Mirjalili et al., 2014). A modified particle swarm optimization and cuckoo search algorithm have been applied to optimize the nonlinear reliability function of a complex system (Kumar et al., 2017; Pant et al., 2017a). Kumar et al. (2019a) have also implemented the multi-objective grey wolf optimization algorithm to optimize the reliability cost functions of the residual heat removal system of a nuclear power plant safety system and the life support system in a space capsule (Kumar et al., 2019b). These algorithms have been used to solve the system of nonlinear equations (Kumar et al., 2018; Pant et al., 2019). Additionally, the literature demonstrates that many nature-inspired algorithms have also been implemented to compute the parameters and optimal value of complex nonlinear functions (Negi et al., 2021; Pant et al., 2017b; Uniyal et al., 2020).

The study optimizes a nonlinear error function to estimate the model parameters using MLE to enhance the forecasting precision. The key reasons prompt to derive the modified model. First, MLP predicts the sign of residuals with high precision, which is required if the observed data sequence has few elements. Second, MLE is used to estimate the model parameters through error optimization. MLE does not involve any assumption, while the formulation of original GM(1,1) assumes background variables in parameters estimation. The empirical study is performed to demonstrate the model validation and forecasting precision. The behaviour of three statistical indicators is assessed to verify the model validation. A comparative analysis is shown among the linear regression, original grey model, and modified grey model based on the forecasting errors.

The grey system theory has been implemented in different areas, namely agriculture (Zeng et al., 2020), medicine (Shen et al., 2019), management (Xuemei et al., 2019), environment (Javed et al., 2020), decision-making (Xuemei et al., 2019), energy (Li & Zhang, 2021), business and economy (Hu et al., 2020). It majorly encompasses grey sequence generation, grey theory and mathematics, relational analysis, forecasting, model building, decision-making, and system control (Lee & Tong, 2011). Generally, GM(1,1) model is implemented on real-world data and generates promising results with a small data set (4 or more data points). This model yields the most reliable results for non-negative data sequences of the exponential form (Wu et al., 2019). The amount of MSW is increasing exponentially around the globe, including Saudi Arabia. The factors of this tremendous growth are rapid increment in population, socio-economic development, fast industrialization, urbanization, and steady improvement in living standards (Araiza-Aguilar et al., 2020). Numerous studies have been published on improved GM(1,1) models with enhanced precision. Hu & Jiang (2017) have successfully applied an enhanced GM(1,1) model on actual data of power demand in China. Hu and Jiang’s model amalgamates the residuals in predicted series and signs of residuals have been predicted using an artificial neural network (ANN) (Hu & Jiang, 2017). A similar approach has also been implemented by Hsu & Chen (2003). Li and Zhang have applied the interpolation method to modify the original data sequence and optimize the background parameters (Li & Zhang, 2018). Hsu and Wang have used Bayesian analysis and genetic algorithms to compute the grey model parameters (Hsu & Wang, 2007; Wang & Hsu, 2008). These models have also been adopted to predict the data in the advanced technology industry, such as integrated circuit manufacturing. A neural network-based GM(1,1) model incorporating genetic programming has been introduced to predict the oil and energy consumption of China (Hu & Jiang, 2017; Lee & Tong, 2011; Yang et al., 2016). Intharathirat et al. (2015) has performed grey relational analysis to determine and rank the waste generation influencing factors. The Univariate and multivariate grey models are derived to predict the amount of waste. The performance comparison has displayed that the univariate grey model has predicted better than naïve and trend curve analysis. Additionally, the multivariate grey model, called the grey model with convolution integral, has considered the most significant influencing factor as input to predict the waste generation amount and has exhibited most forecasting precision than the univariate grey model (Intharathirat et al., 2015). A similar multivariate grey model is implemented to predict electronic waste generated from mobile, computer, and television (Duman et al., 2019; Kiran et al., 2021).

The content of the article is structured into seven sections. Section 1 concentrates on the grey theory of forecasting and its applications both in theoretical and practical domains, and section 2 introduces linear regression and derivation of original GM(1,1). Section 3 illustrates the modelling methodology, consisting of the residual model, model parameters estimation, and MLP for residual signs prediction. Section 4 describes three statistical indicators to validate the model and compare the accuracy, and section 5 illustrates the empirical study. Section 6 analyses and compares the forecasted results of the implemented model with the original grey model and linear regression along with model validation. Section 7 discusses the outcomes and presents the conclusion along with future work.

2. Linear Regression and Original GM(1,1) Model

2.1 Linear Regression

The linear regression model is the most prevalent statistical method to analyse the linear relationship between two or more variables. A simple linear regression model of one independent or explanatory variable, x and one dependent or estimated variable, y is defined by the following linear equation:

$$y = \beta x + \alpha \tag{1}$$

where, α is a constant or intercept, while β is called regression coefficient or slope. α and β are calculated by the least square method.

2.2 GM(1,1) Model

The GM(1,1) model belongs to the category of homogeneous exponential growth models to predict time series. The basic operations in modelling involve computation of cumulative sequence, inverse cumulative operation, formation of the grey differential equation, and the estimation of parameters, namely developing coefficient and control variable. The precision of forecasting is substantially influenced by the growing trends of a raw data sequence, and a smooth data sequence with nearly exponential growth deliver more precise results. The GM(1,1) model construction needs more than four consecutive data items with equal intervals and a differential equation to describe the original data sequence. Therefore, few data items are sufficient to characterize the original raw data sequence and model parameter estimation. This section illustrates the GM(1,1) modelling basic operations, namely: (1) Accumulated Generating Operation (AGO): Most consequential characteristics of the grey system theory that recognizes the hidden consistency in the actual sequence to increase the prediction precision, decrease the arbitrary features of data series and produce new cumulative data sequence (2) Inverse Accumulated Generating Operation (IAGO): It utilizes difference method to reconstruct the monotonic sequence and obtain predicted sequence (3) Grey Modelling: It comprises the formation of a differential equation, and it's parameter estimation. Following are the steps involved in building the original GM(1,1) model.

Step I: Raw Data Representation in Time Series

The original non-negative raw data sequence provided by any system is considered as follow:

$$W^{(0)} = \left\{ w^{(0)}(1), w^{(0)}(2), \dots, w^{(0)}(n) \right\}, n \geq 4,$$

where, $w^{(0)}(k)$, $k = 1, 2, \dots, n$; is the data item in the time series at period k, and n represents the total number of observed periods.

Step II: Computation of AGO and Formation of the Grey Differential Equation

AGO converts the grey system into the white system and identify development trend in grey data. Perform the AGO function at $W^{(0)}$ to generate the new sequence, which is monotonically increasing, as follows:

$$W^{(1)} = \left\{ w^{(1)}(1), w^{(1)}(2), \dots, w^{(1)}(n) \right\}.$$

where, $w^{(1)}(k)$ is computed from $w^{(1)}(k) = \sum_{j=1}^{k} w^{(0)}(j)$, $k = 1, 2, \dots, n$.

The above monotonic series is represented by the following grey differential equation:

$$\frac{dw^{(1)}(k)}{dk} + uw^{(1)}(k) = v \tag{2}$$

u and v are coefficients of the grey model, and u is known as developing coefficient, and v is called endogenous control variable or the grey input, and entirely depend on background information. The solution of the above differential equation gives the predicted value $\hat{w}^{(1)}(k)$ with $\hat{w}^{(1)}(1) = w^{(0)}(1)$. After solving equation (2), the expression of $\hat{w}^{(1)}(k)$ is given as:

$$\hat{w}^{(1)}(k)=\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(k-1)\right)+\frac{v}{u},\ k=2,3,\ldots,n.$$

Step III: Parameter Estimation

The least-square method is applied to estimate u , and v . Assume parameter sequence $\hat{a}$, $\hat{a}=[u,v]^T$ and $\hat{a}$ can be computed using the following formula as follows:

$$\hat{a}=[u,v]^T=\left(A^TA\right)^{-1}A^TB \tag{3}$$

where, $A=\begin{bmatrix}-\omega^{(1)}(2) & 1\\ -\omega^{(1)}(3) & 1\\ \vdots & \vdots\\ -\omega^{(1)}(n) & 1\end{bmatrix}$ $\omega^{(1)}(k)=\lambda w^{(1)}(k-1)+(1-\lambda)w^{(1)}(k),\ k=2,3,\ldots,n.$

$B=\left[w^{(0)}(2),w^{(0)}(3),\ldots,w^{(0)}(n)\right]^T.$

where, A and B are data array and data column respectively, $\omega^{(1)}(k)$ is intermediate computational value. Generally, the value of λ is assumed as 0.5 for simplification, but this assumption is not always optimal. Therefore, u and v entirely depend on $\omega^{(1)}(k)$, which is extremely difficult to calculate.

Step IV: Computation of IAGO and Generation of Predicted Sequence

Perform IAGO to revert the AGO and reconstruct the monotonic sequence to find the forecast value $\hat{w}^{(0)}(k)$.

$$\hat{w}^{(0)}(k)=\hat{w}^{(1)}(k)-\hat{w}^{(1)}(k-1).$$

Therefore,

$$\hat{w}^{(0)}(k)=\left(1-\exp(u)\right)\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(k-1)\right),\ k=2,3,\ldots,n \tag{4}$$

where, $\hat{w}^{(0)}(1)=w^{(0)}(1)$ holds.

Now, the forecasted series will be $\hat{W}^{(0)}=\left\{\hat{w}^{(0)}(1),\hat{w}^{(0)}(2),\ldots,\hat{w}^{(0)}(n)\right\}$.

3. Modelling Methodology

A modified grey forecasting model is derived and applied to the actual data of MSW generation to uncover the forecasting power. Various statistical indicators are computed to perform the model validation, and forecasted results are compared with the GM(1,1) model and linear regression.

Modelling methodology involves five steps as follows: (i) Derivation of original grey model (ii) Construction of residual grey model (iii) Computation of model parameters using maximum likelihood estimation (MLE) (iv) Residual sign estimation using MLP (v) Finally, the combination of the original and residual model along with signs to derive the modified model.

This research embodies MLP based estimated signed residuals with the original GM(1,1) to construct a more powerful forecasting model. The MLP has the potential to predict the signs of the residuals more accurately. Therefore, integration of residuals significantly increases the forecasting precision, which is extensively demanded in the future estimation of MSW generation. MSWM planning and infrastructure development depend on forecasted MSW generation. The procedure of modified model building is outlined in Figure 1. The following subsections explain the detailed process of deriving the modified grey model.

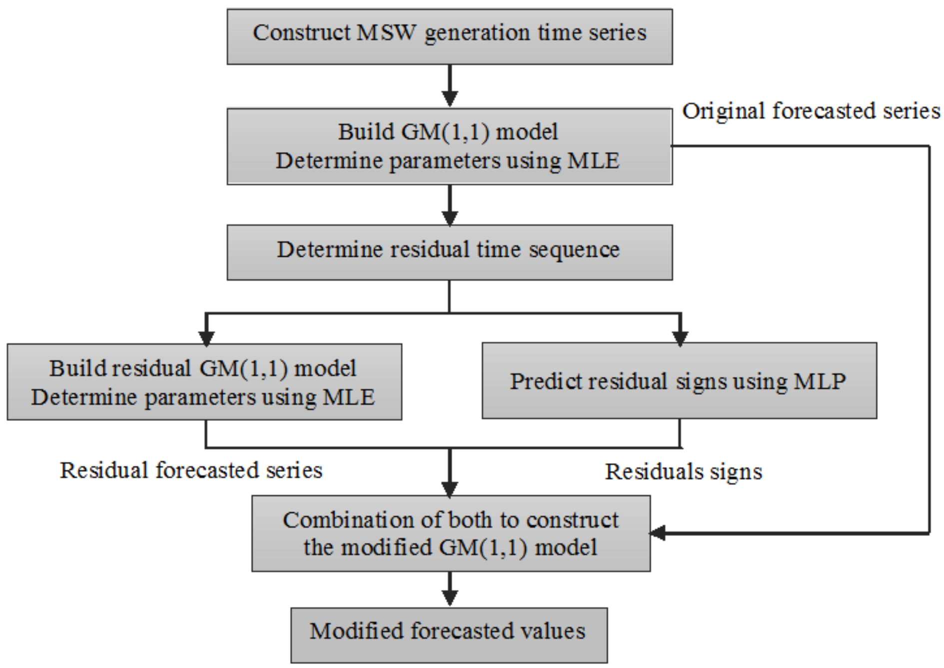

Figure 1. The procedure of developing a modified grey model.

3.1 Residual GM(1,1) Formulation

The original grey model is applied to the residual sequence to determine forecasted residuals. Residual series is a sequence of differences of corresponding values between the original series $W^{(0)}$ and forecasted series $\hat{W}^{(0)}$. A sub-model is constructed that contains two primary operations, namely residual series forecasting and sign estimation. The original GM(1,1) model is applied on absolute residual series $E^{(0)}$ to obtain the predicted residual sequence $\hat{E}^{(0)}$, and the MLP technique is used to determine the signs of forecasted residuals. Then, residuals with the computed signs are combined with the original forecasted series to obtain the final modified predicted sequence. The computational algorithm to implement the sub-model is explained as follows.

Step I: Generate Forecasted Series
Construct the forecasted sequence $\hat{W}^{(0)}$.

Step II: Generate Residual Series
Yield absolute residual series value as follows.

$$E^{(0)} = \left\{ e^{(0)}(2), e^{(0)}(3), ..., e^{(0)}(n) \right\}.$$

where, $e^{(0)}(k) = \left| w^{(0)}(k) - \hat{w}^{(0)}(k) \right|, \ k = 2, 3, \ldots, n$.

Step III: Construct Residual Model
Now, forecasted residual $\hat{e}^{(0)}(k)$ can be given in similar ways of $\hat{w}^{(0)}(k)$, as:

$$\hat{e}^{(0)}(k) = \left(1 - \exp(u_e)\right)\left(e^{(0)}(2) - \frac{v_e}{u_e} \right)\exp\left(-u_e(k-1)\right), \ k = 3, 4, \ldots, n \tag{5}$$

where, initial condition $\hat{e}^{(0)}(2) = e^{(0)}(2)$ holds and u_e and v_e is termed as developing coefficient and control variable for residual series, and entirely depend on the intermediate information.

The forecasted residual series will be $\hat{E}^{(0)} = \left\{ \hat{e}^{(0)}(2), \hat{e}^{(0)}(3), ..., \hat{e}^{(0)}(n) \right\}$.

Step IV: Perform Integration of Residuals
The improved forecasted value $\hat{w}'^{(0)}(k)$ is determined by summing up signed $\hat{e}^{(0)}(k)$ with $\hat{w}^{(0)}(k)$. The expression of $\hat{w}'^{(0)}(k)$ is shown as follows:

$$\hat{w}'^{(0)}(k) = \hat{w}^{(0)}(k) + s(k)\hat{e}^{(0)}(k), \ k = 1, 2, \ldots, n \tag{6}$$

where, $s(k)$ represents the sign of residual $\hat{e}(k)$. Generally, $s(k)$ depends on the method used for sign estimation. Lee and Tong (2011) have implemented a genetic program to determine the sign of residual. While Hsu and Chen (2003), and Hu and Jiang (2017) have proposed an ANN based algorithm to estimate the signs. In this paper, the computation of $s(k)$ is performed through the proposed MLP based method. The detailed formulation of this method is explained in the next section.

3.2 MLP Based Residual Sign Estimation
Artificial intelligence (AI) techniques have emerged as the most powerful computational tool to solve time series forecasting problems in the recent decades. Much of the literature published contains a massive number of applications of different AI techniques. Among the vast number of AI techniques, ANNs are the most common method applied in time series forecasting (Lin et al., 2018). ANNs are analogous to real biological neural networks but in uncomplicated form and designed to simulate the operations of real brains (Tang et al., 2019; Zhang et al., 2019). A simple

biological neural model and corresponding computational neural model architecture are shown in Figure 2. The fundamental element of processing in ANNs is a neuron, also called a node. It takes the input signal from subsequent neurons or external sources and performs some processing to generate output (Elsheikh et al., 2019). Each input has a certain level of significance for other inputs; therefore, distinct weights are assigned to each input. Then, a function $\varphi(\bullet)$ is computed on a weighted sum of inputs, as shown in Figure 2. This function is called the activation function, and it may be linear and nonlinear. The objective of the nonlinear activation function is to produce nonlinearity in output. It is fundamentally essential as a large volume of real-world data are nonlinear; therefore, it is required that neurons must be able to learn nonlinearities to solve realistic problems. Each activation function has a particular type of mathematical expression and receives a single value for computation. The mathematical expression of output Y at any node is shown in the following equation.

$$Y = \varphi(\sum_{i=1}^{m} w_{ki} * x_i + b) .$$

There are numerous activation functions in practice, namely logistic, hyperbolic tangent, rectified linear unit (ReLU), and other nonlinear functions. The logistic activation function is used to implement MLP and has the following mathematical expression.

$$\varphi(x) = \frac{e^x}{1+e^x}, x \in R .$$

where, $\varphi(\bullet)$ lies between 0 and 1.

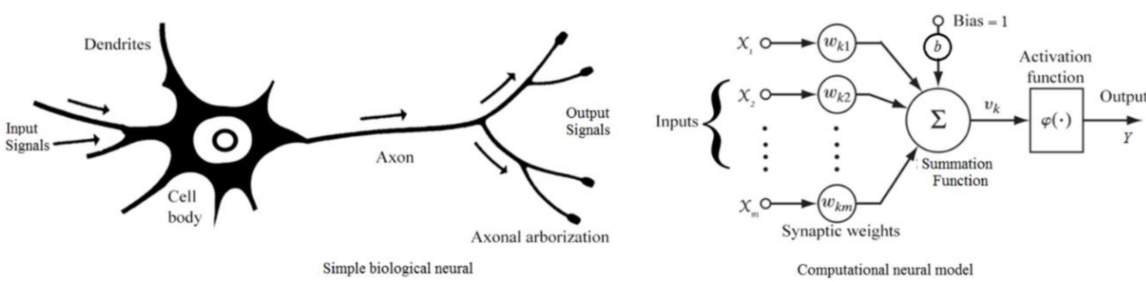

Figure 2. Simple biological neural versus computational neural model.

The objective of ANNs building is to develop computer systems with adaptive learning and generalized processing capabilities like human brains. They can solve large and complex systems having numerous interconnected parameters without requiring the characteristics of the system. They can be trained to detect similar types of inputs and generate a specific output corresponding to that input. ANNs have numerous types of neural network architectures to solve time series forecasting problems. Among these architectures, most used is MLP which is also called a multilayer feed-forward neural network and consists of three or more perceptron (neuron) layers: input layer, an output layer, and a minimum one hidden layer (Sun & Huang, 2020). Neurons in one layer are connected to another layer only in the forward direction. The general architecture of the MLP is depicted in Figure 3 (Heidari et al., 2019). The accuracy of the ANNs model mainly depends on the training of the neural network as in this phase network iteratively adjusts initially assigned weights to learn the relationship between input and output. Many training algorithms exist to train the MLP, but the backpropagation algorithm (BPA) is the most prevailing among these. The basic principle of BPA permits propagating the error in the backward direction, which facilitates the improvement of the weights. This backpropagation is performed iteratively to minimize the loss function, which is the difference between the actual value and the model predicted value. After each iteration, weights approach optimal values, giving minimum loss and providing the most precise predicted output. In BPA implementation, the gradient descent (GD) technique minimizes the loss function. The GD algorithm searches the local or global minima of the loss function iteratively. The loss function in MLP with backpropagation is the mean squared error E shown in the following equation (Qu et al., 2019).

$$E = \frac{1}{2N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2 .$$

where, y_i is the actual value corresponding to input x_i while $\hat{y}_i$ is calculated value for the input x_i .

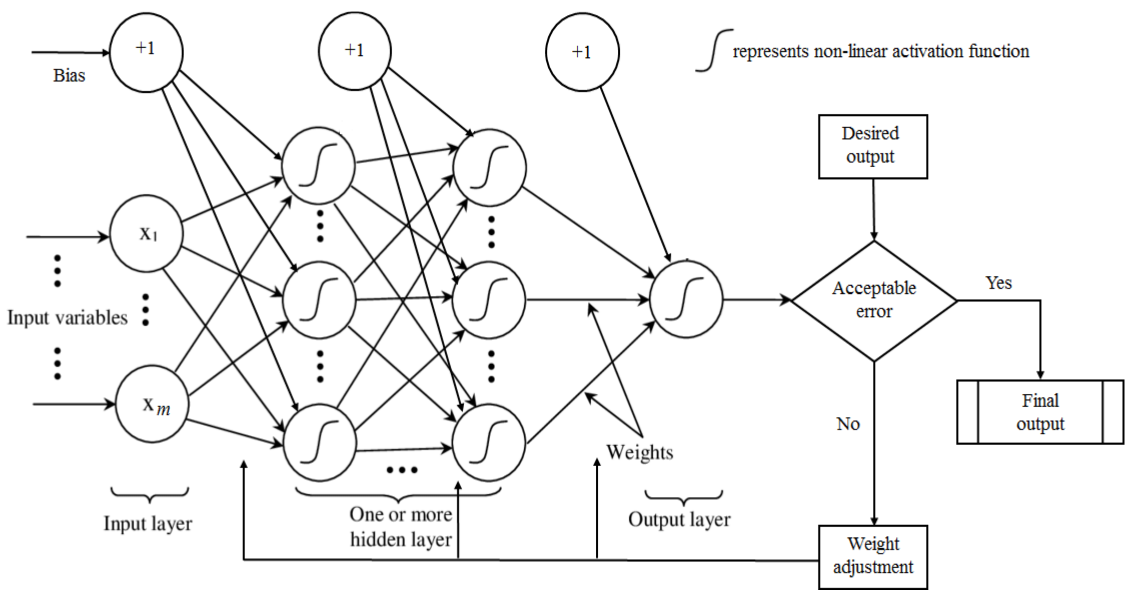

Figure 3. A typical MLP network with backpropagation.

A two-input with one hidden layer MLP network is designed to estimate the signs of predicted residual sequence. For model implementation, a variable $r(k)$ is assumed to simulate the sign of k^{th} residual. If the k^{th} residual has a plus sign, then $r(k)$ will have value +1, and in all other conditions, it will be 0. Now, consider $s(k)$ represents the sign of k^{th} residual which is defined as follows:

$$s(k) = \begin{cases} +1; \text{ if } r(k) = 1 \\ -1; \text{ if } r(k) = 0 \end{cases} ; k = 1, 2, \ldots, n .$$

Now, from equation (4), (5), and (6), the modified forecasted value $\hat{x}'^{(0)}(k)$ is derived below:

$$\hat{w}'^{(0)}(k) = \left(1-\exp(u)\right)\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(k-1)\right) + s(k)\left(1-\exp(u_e)\right)\left(e^{(0)}(2)-\frac{v_e}{u_e}\right)\exp\left(-u_e(k-1)\right) \quad (7)$$

where, $k=1,2,\ldots,n$.

3.3 Model Parameter Estimation: MLE Method

Since parameters u and v depend entirely on the background value $\omega^{(1)}(k)$, which is very tedious to calculate. Therefore, the MLE approach is proposed to compute model parameters. This approach is entirely independent from the computation of $\omega^{(1)}(k)$. The following equation is established to implement the estimation of the MLE-based parameters.

Assume an independent variable t_i, indicates year number, i.e., k and dependent variable y_i, which represents the amount of waste generated in that year, now the following equation will always hold.

$$y_i = \left(1-\exp(u)\right)\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(t_i-1)\right)+\varepsilon_i;\ i=1,2,\ldots,n$$

$$\varepsilon_i = y_i - \left(1-\exp(u)\right)\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(t_i-1)\right) \quad (8)$$

Since MLE estimates the parameters of a probability distribution function (pdf), therefore, it is considered that residuals $e(i)$ indicate errors ε_i and are assumed as random variables which are independent and identically distributed and follow the normal distribution with mean μ and variance σ^2. The pdf of the normal distribution $\varepsilon_i \sim N(\mu,\sigma^2)$ is given in equation (9).

$$f\left(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n \mid \mu,\sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_i-\mu}{\sigma}\right)^2\right) \quad (9)$$

Now, substitute ε_i from equation (8) in the above equation

$$f\left(t_1,t_2,\ldots,t_n \mid \mu,\sigma^2,u,v\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\left(y_i-\left(1-\exp(u)\right)\left(w^{(0)}(1)-\frac{v}{u}\right)\exp\left(-u(t_i-1)\right)-\mu\right)^2}.$$

Assume, $\mu=\theta_0$, $\sigma^2=\theta_1$, $u=\theta_2$ and $v=\theta_3$. Now, the above equation will be:

$$f\left(t_1,t_2,\ldots,t_n \mid \theta_0,\theta_1,\theta_2,\theta_3\right)=\frac{1}{\sqrt{2\pi\theta_1}}e^{-\frac{1}{2\theta_1}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right)^2} \tag{10}$$

Famous British statistician Sir Ronald Aylmer Fisher proposed that the likelihood function is pdf of the data, which is the function of parameters. The likelihood function computes the goodness of fit of the statistical model to an observed data set for a set of values of unknown model parameters. So, values of parameters are estimated for maximum fitness of model to observed data, which means maximization of the likelihood function. Now, the likelihood function L of random variables ε_i can be represented as:

$$L=\prod_{i=1}^{n} f\left(t_1,t_2,\ldots,t_n \mid \theta_0,\theta_1,\theta_2,\theta_3\right).$$

Now, putting function value from equation (10) in the above equation.

$$L=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\theta_1}}\exp\left(-\frac{1}{2\theta_1}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right)^2\right).$$

The log-likelihood equation of L is given by:

$$LL=\ln L=\sum_{i=1}^{n}\left(-\ln\sqrt{2\pi\theta_1}-\frac{1}{2\theta_1}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right)^2\right)$$

$$LL=\ln L=-\sum_{i=1}^{n}\ln\sqrt{2\pi\theta_1}-\sum_{i=1}^{n}\frac{1}{2\theta_1}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right)^2 \tag{11}$$

MLEs of $\theta_0,\theta_1,\theta_2$ and θ_3 is obtained by solving equations $\frac{\partial LL}{\partial\theta_0}=0$, $\frac{\partial LL}{\partial\theta_1}=0$, $\frac{\partial LL}{\partial\theta_2}=0$, and $\frac{\partial LL}{\partial\theta_3}=0$.

$$\frac{\partial LL}{\partial\theta_0}=\sum_{i=1}^{n}\frac{1}{\theta_1}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right) \tag{12}$$

$$\frac{\partial LL}{\partial\theta_1}=-\sum_{i=1}^{n}\frac{1}{2\theta_1}+\sum_{i=1}^{n}\frac{1}{2\theta_1^2}\left(y_i-\left(1-\exp\left(\theta_2\right)\right)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp\left(-\theta_2\left(t_i-1\right)\right)-\theta_0\right)^2 \tag{13}$$

$$\frac{\partial LL}{\partial\theta_2}=\sum_{i=1}^{n}\frac{1}{\theta_1}\left(\begin{array}{r}\left((1-t_i)\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)+\frac{\theta_3}{\theta_2{}^2}\right)(1-\exp(\theta_2))\exp(-\theta_2(t_i-1))\\ -\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp(-\theta_2(t_i-2))\end{array}\right)$$
$$*\left(y_i-(1-\exp(\theta_2))\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp(-\theta_2(t_i-1))-\theta_0\right) \tag{14}$$

$$\frac{\partial LL}{\partial\theta_3}=-\sum_{i=1}^{n}\frac{1}{\theta_1\theta_2}\left(y_i-(1-\exp(\theta_2))\left(w^{(0)}(1)-\frac{\theta_3}{\theta_2}\right)\exp(-\theta_2(t_i-1))-\theta_0\right)$$
$$*(1-\exp(\theta_2))\exp(-\theta_2(t_i-1)) \tag{15}$$

MLEs of θ are the values of θ that maximize the likelihood function L or log-likelihood function LL. For these values of θ , i.e., set of model parameters, the model is most probable to predict the data most close to the observed data.

4. Indicators of Forecasting Model Validation

The implemented models must be investigated according to certain indicators and must satisfy the criteria before forecasting. In statistics, various indicators can be employed to examine the model validation. Three indicators have been implemented to assess validation performance; these are relative residual, absolute degree of grey indices, and posterior variance test. The computation of each indicator is given as follows:

4.1 Relative Residual

Relative residual is percentage error (PE), in which predicted values are compared with corresponding actual values to examine the relative errors whether they meet the requirement. The relative residue of the grey model is $\phi(i)$, as the formula:

$$\phi(i)=\frac{\chi(i)}{w^0(i)}=\frac{\hat{w}^{(0)}(i)-w^{(0)}(i)}{w^{(0)}(i)}\times 100\%\text{, where, }\chi(i)=\hat{w}^{(0)}(i)-w^{(0)}(i) \tag{16}$$

The mean absolute relative residual or mean absolute percentage error (MAPE) $\bar{\phi}$ is given as:

$$\bar{\phi}=\frac{1}{n}\sum_{i=1}^{n}|\phi(i)| \tag{17}$$

The accuracy of the model is p^0 as the formula:

$$p^0=(100-\phi).$$

The proposed model can be categorized into four discrete classes of forecasting accuracy depending on the value of p^0 (Hu & Jiang, 2017). The value of p^0 and corresponding accuracy category are shown in Table 1.

Table 1. The forecasting accuracy grade table.

p^0	[90%, 100%]	[80%, 90%)	[50%, 80%)	[0%, 50%)
Forecasting accuracy	High	Good	Reasonable	Weak
$p^0 \geq 80\%$ is the requirement of model validation.				

4.2 Posterior Variance Test

The posterior variance test is a statistical method based on the evaluation of variance and probability distribution of residuals. Two pivotal indexes of precision, namely posterior variance-ratio C and posterior probability P are computed to quantify the class of the model. The condition of high forecasting accuracy, C should be as minimum as possible and P should have as maximum as possible value. According to C and P values, the accuracy of the model can be quantified into four distinct forecasting accuracy classes, namely "high, good, reasonable, and weak" (Wang et al., 2020). The posterior variance ratio C is computed as:

$$C = \frac{S_2}{S_1} \tag{18}$$

where, $S_1{}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(w^{(0)}(i) - \bar{w}^{(0)}\right)^2$ and $S_2{}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\chi(i) - \bar{\chi}\right)^2$.

Average of the original sequence, $\bar{x}^{(0)}$ and residual series, $\bar{\chi}$ are formula as:

$$\bar{w}^{(0)} = \frac{1}{n}\sum_{i=1}^{n} w^{(0)}(i) \text{ and } \bar{\chi} = \frac{1}{n}\sum_{i=1}^{n}\left|\chi(i)\right|.$$

The posterior probability P is given as:

$$P = P\left\{\left|\chi(i) - \bar{\chi}\right| < 0.6745 S_1\right\} \tag{19}$$

Table 2. The forecasting accuracy quantification table.

C	< 0.35	< 0.45	< 0.50	≥ 0.65
P	> 0.95	> 0.80	> 0.70	≤ 0.70
Forecasting accuracy	High	Good	Reasonable	Weak

4.3 Test of Absolute Degree of Grey Indices

It is also called the correlation degree test, which determines the degree of similarity between observed and forecasted curves. Based on general intuition, more closeness between curves implies more similarity in change pattern means high correlation. The correlation degree is ξ as the formula:

$$\xi = \sum_{i=1}^{n} \xi(i).$$

where, $\xi(i)$ is correlation coefficient of $\hat{W}^{(0)}$ and $W^{(0)}$ as the formula:

$$\xi(i) = \frac{\min_{1\le i\le n}|\chi(i)| + \rho \max_{1\le i\le n}|\chi(i)|}{|\chi(i)| + \rho \max_{1\le i\le n}|\chi(i)|} \tag{20}$$

where, $\rho \in [0, 1]$ is called resolution ratio or distinguishing coefficient $\rho = 0.5$. In general, ξ must be greater than 0.6 for high correlation, which implies better forecasting (Ren et al., 2020).

5. Empirical Study

Cleanliness is one of the essential factors of sustainable urban development to maintain a clean environment and protecting resident's health. MSW is one of the major sources that significantly affect cleanliness. Therefore, MSWM has emerged as a crucial concern to the local governments. The MSW planning and demand of resources depend substantially on the amount of waste that will be managed. Accurate forecasting is a mandatory prerequisite to developing an effective MSWM plan, as it can provide the optimal number of resources. Therefore, finding the precise prediction of MSW generation is very important. A real case of MSW generation in Saudi Arabia is studied to illustrate the precision of the modified grey forecasting model. The historical annual municipal solid waste generation from 2010 to 2016 is considered to estimate the unknown parameters of the model, i.e., training data set and data of the year 2017 and 2018 are employed as ex-post testing. After computation of parameters, the equations of simple linear regression, original GM(1,1) and modified GM(1, 1) models are given as:

***GM(1,1) Model*:** It is applied to the studied waste generation time series to obtain the forecasting equation. The parameters a and b are estimated using equation (3) and put in equation (4), then equation (21) is derived, as follows:

$$\hat{x}^{(0)}(k) = 11362.691 * \exp(0.0672025137 * (k-1)), \; k = 2, 3, \ldots, n \tag{21}$$

***Modified GM(1,1) Model*:** The parameters a and b of both models are estimated using MLE equations (12), (13), (14), and (15). The estimates of a and b are put in equation (7) and equation (22) is shown as:

$$\hat{x}^{(0)}(k) = 9607.054 * \exp(0.1024146589 * (k-1)) + s(k) * 454.004 * \exp(0.2018948811 * (k-1)), \; k = 1, 2, \ldots, n \tag{22}$$

The sign $s(k)$ is predicted using the MLP as explained in section 4.2.2.

Linear Regression: The linear regression equation coefficients α and β are estimated using the least square method, and coefficient estimates are put in equation (3) to get equation (23), as follows:

$$y = 643.61 * x + 10693.29 \quad (23)$$

6. Result and Discussion

According to the studied literature, it can be concluded that various methods can be used to predict the amount of MSW generation. These methods can be broadly categorized into three categories: time series forecasting, multivariate analysis, and AI models. In the above prediction methods, the fundamental factor that significantly affects the forecasting capability of the model is the training data size, which defines the significance of their predictions. The available MSW generation data is limited in size and generally exhibits an exponential trend. Therefore, all these methods with limited data and exponential trends are not appropriate. Then, in this condition, it is suitable to construct the grey model to determine the MSW generation forecasting. In the study, the original GM(1,1) and modified GM(1,1) models are derived, and linear regression is also considered to compare the implemented model performance. The empirical study is performed, and the amount of MSW is predicted using the above three models. Before evaluating the modified GM(1,1) model performance and explaining the significance of results, the validation of two models is performed as follows.

The original and modified grey model validation is performed based on three criteria demonstrated in section 4. Table 3 summarizes the performance of statistical indicators for both original and modified models' validation. The values of relative residual criterion p^0 of the original GM(1,1) and modified GM(1,1) models are 94.39% and 97.89%, respectively. According to Table 2, it is concluded that both models lie in the high category of forecasting accuracy, as their p^0 values lie in the interval of (90%, 100%). Additionally, the criteria p^0 clearly shows that the modified model exhibits high precision compared to the original GM(1,1). For the original GM(1,1) model, the values of the test of the absolute degree of grey indices are C = 0.44, P = 0.86, which lie in the good forecasting accuracy, and the value of post variance test ξ = 0.6 also lies the in the good forecasting accuracy. Similarly, for the modified GM(1,1) model, the values of the test of the absolute degree of grey indices C = 0.22, P = 1.00 and post variance test ξ = 0.67 belong to the high forecasting accuracy. Finally, the test of the absolute degree of grey indices and post variance test validate the grey and modified grey models with good and high forecasting accuracy, respectively. Therefore, it can be concluded that both models are valid and reliable, but the modified model displays greater prediction accuracy.

Table 3. The comparison of statistical indicators.

Validation Indicator	GM(1,1)		Modified GM(1,1)	
	Value	Forecasting accuracy	Value	Forecasting accuracy
p^0	5.61	High	2.19	High
C P	0.44 0.86	Good	0.22 1.00	High
ξ	0.60	Good	0.67	High

The study considers the linear regression, original GM(1,1) and modified GM(1,1) models to simulate and predict the amount of MSW generation in Saudi Arabia from 2010 to 2018. The simulative and forecasted data, along with the actual data, are displayed in Table 4. These data represent that the MSW generation in Saudi Arabia is nonuniform. It has increased from 2010 to 2017, while it has decreased from 2017 to 2018. Table 4 and Figure 4 present the forecasted results of the applied models and compare them with actual values. Figure 5 depicts the distribution of fitting and forecasting PE of all models. In Table 4, the MAPE of linear regression, GM(1,1) and modified model for model fitting data (2010-2016) are and 1.73% , 5.61%, and 2.19% respectively and for test data set (2017-2018) are 27.90%, 15.97%, and 8.90% respectively. The modified GM(1,1) model is more powerful in forecasting the MSW generation amount the original GM(1,1) and linear regression, as it has the lowest prediction MAPE among all models, while linear regression is slightly more powerful in simulation as compared to the modified GM(1,1) model. This result exhibits that the modified grey model has exceptional performance.

The ex-post testing accuracy analysis indicates that the modified GM(1,1) model exhibits more precision of prediction than other models. It has a lower error of 8.90%, compared to errors of 15.97% and 27.90%. One remarkable observation can be seen in the results that the linear regression model exhibits slightly higher prediction accuracy than the other two models in the training data phase. The linear regression errors, original and modified grey models in training are 1.73%, 5.61% and 2.19%, respectively. The errors of the three models in ex post-testing are 27.90%, 15.97% and 8.90%. These values imply that the modified model demonstrates a substantially higher accuracy than the original GM(1,1) and linear regression in this phase. Therefore, it can be confidently deduced that the modified model is superior to the other two models in terms of future forecasting. The key criteria for evaluating a forecasting model should be more concentrated on generalization ability than model fitting. The Linear regression model experiences over-fitting as the regression line is very close to the original curve in the training phase while very far ex post-testing, shown in Figure 4. Therefore, linear regression cannot be generalized for good forecasting results. In contrast, the modified grey model has more capability of generalization than the original grey model, as the modified model exhibits lower errors than the original GM(1,1), depicted in Figure 5. Finally, it can be concluded that the modification approach significantly increases the performance and reliability of the original grey model.

Table 4. Comparison of forecasted values and performance measurement of different models.

Year	**Actual data** (Unit: 10^3 tons)	**GM(1,1)**		**Modified GM(1,1)**		**Linear Regression**	
		Forecasted	**Error***	**Forecasted**	**Error***	**Forecasted**	**Error***
2010	11,555	11,555.00	0.00	11555.00	0.00	11336.90	-1.89
2011	12,048	12,152.53	0.87	12048.00	0.00	11980.51	-0.56
2012	12,560	12,997.28	3.48	12470.75	-0.71	12624.12	0.51
2013	13,093	13,900.75	6.17	13894.41	6.12	13267.73	1.33
2014	13,646	14,867.02	8.95	13453.05	-1.41	13911.34	1.94
2015	14,220	15,900.46	11.82	14785.89	3.98	14554.95	2.36
2016	15,752	17,005.73	7.96	16236.07	3.07	15198.56	-3.51
MAPE# (%) (2010-2016)			5.61		2.19		1.73
2017	24,213	18,187.83	-24.88	21541.67	-11.03	15842.17	-34.57
2018	20,930	19,452.11	-7.06	19514.88	-6.76	16485.78	-21.23
MAPE# (%) (2017-2018)			15.97		8.90		27.90

*Refer to equation (16); # Refer to equation (17); Data Source: (General Authority for Statistics, 2018).

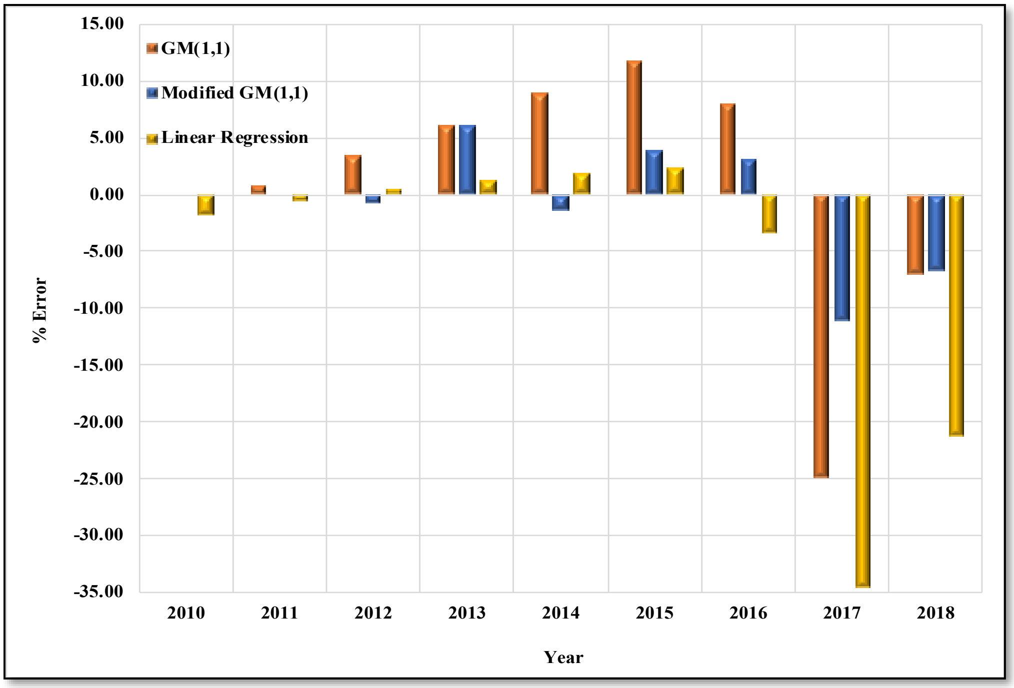

Figure 4. Trends of forecasted and target values.

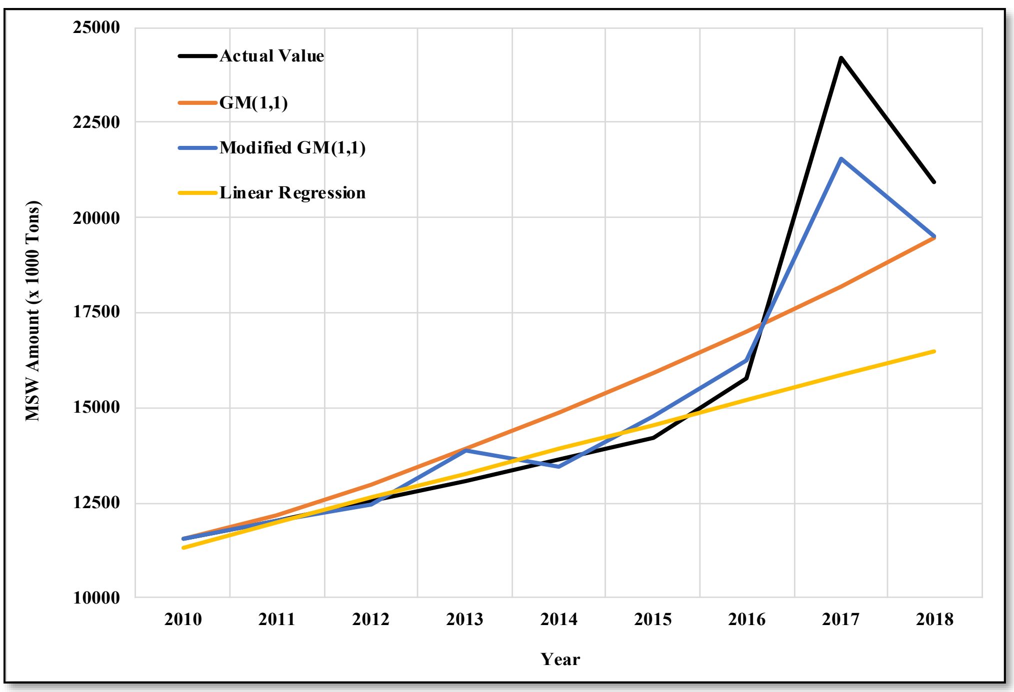

Figure 5. Distribution of forecasting percentage errors.

7. Conclusions

Accurate forecasting of MSW generation is very complex since it depends on various influencing factors such as population growth, urban settlement, economic development, and increasing living standards. Therefore, an accurate forecasting model is demanded for better MSWM planning and accurate resource prediction. The GM(1,1) model is modified by incorporating residuals through MLP based sign prediction to forecast the MSW generation precisely. The modified model is applied to actual data of MSW generation and compared with the original grey model and linear regression. The goodness-of-fit tests verify that the precision of modified GM(1,1) is qualified for accuracy class "high" (C = 0.22, P = 1.00, ξ = 0.60), while the GM(1,1) belongs to class "good" (C = 0.44, P = 0.86, ξ = 0.60). The linear regression model experiences over-fitting as forecasting accuracy is very high in training data p^0 = 1.73%, while very low p^0 = 27.90% for testing data. The MAPEs of all models is rigorously examined and compared, and then it is obtained that the GM(1,1) forecasting precisions, the modified GM(1,1), and linear regression is 15.97%, 8.90%, and 27.90%, respectively. The modified grey model achieves the lowest MAPE in post forecasting; therefore, it can be stated that the residual correction approach can substantially increase the precision of forecasted results over the original prediction.

The future work and limitation of the study are given as follows. The implemented model is monotonic exponential in nature; therefore, it is a modelling prerequisite. Various influencing factors such as rapid growth in population, socio-economic development, fast industrialization, urbanization, and improvement in living standards impact the amount of waste generation, but the model does not incorporate these factors as input variables to predict the waste generation amount. A study can be performed to consider the maximum possible influencing factors to produce better forecasting results. Furthermore, the modified GM(1,1) model does not need a large sample size and is less complex than other statistical and AI models. Besides the above, the reliability and forecasting accuracy of the model has also been demonstrated with empirical study, so it can also be utilized in other research areas. Based on the literature analysis, it can be concluded that metaheuristic algorithms are a potential option to optimize the error function for computing the model parameters. These algorithms have been implemented in various studies to compute the parameters and optimal value of nonlinear functions. Nowadays, researchers from different fields are exploiting the power of metaheuristic algorithms to solve their optimization problems.

Conflict of Interest

The authors confirm that there is no conflict of interest to declare for this publication.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

Araiza-Aguilar, J.A., Rojas-Valencia, M.N., & Aguilar-Vera, R.A. (2020). Forecast generation model of municipal solid waste using multiple linear regression. *Global Journal of Environmental Science and Management*, *6*(1), 1-14. DOI: https://doi.org/10.22034/gjesm.2020.01.01.

Balochian, S., & Baloochian, H. (2020). Improving grey prediction model and its application in predicting the number of users of a public road transportation system. *Journal of Intelligent Systems*, *30*(1), 104-114. DOI: https://doi.org/10.1515/jisys-2019-0082.

Duman, G.M., Kongar, E., & Gupta, S.M. (2019). Estimation of electronic waste using optimized multivariate grey models. *Waste Management*, *95*, 241-249. DOI: https://doi.org/10.1016/j.wasman.2019.06.023.

Elsheikh, A.H., Sharshir, S.W., Abd Elaziz, M., Kabeel, A.E., Guilan, W., & Haiou, Z. (2019). Modeling of solar energy systems using artificial neural network: a comprehensive review. *Solar Energy*, *180*, 622-639. DOI: https://doi.org/10.1016/j.solener.2019.01.037.

General Authority for Statistics. (2018). *Per capita daily waste collection In Saudi Arabia during the period 2010-2018*. https://www.stats.gov.sa/sites/default/files/Per capita waste generation 2018 EN.pdf.

Heidari, A.A., Faris, H., Aljarah, I., & Mirjalili, S. (2019). An efficient hybrid multilayer perceptron neural network with grasshopper optimization. *Soft Computing*, *23*(17), 7941-7958. DOI: https://doi.org/10.1007/s00500-018-3424-2.

Hsu, C.C., & Chen, C.Y. (2003). Applications of improved grey prediction model for power demand forecasting. *Energy Conversion and Management*, *44*(14), 2241-2249. DOI: https://doi.org/10.1016/S0196-8904(02)00248-0.

Hsu, L.C., & Wang, C.H. (2007). Forecasting the output of integrated circuit industry using a grey model improved by the bayesian analysis. *Technological Forecasting and Social Change*, *74*(6), 843-853. DOI: https://doi.org/10.1016/j.techfore.2006.02.005.

Hu, Y.C. (2020). Constructing grey prediction models using grey relational analysis and neural networks for magnesium material demand forecasting. *Applied Soft Computing*, *93*. DOI: https://doi.org/10.1016/j.asoc.2020.106398.

Hu, Y.C., & Jiang, P. (2017). Forecasting energy demand using neural-network-based grey residual modification models. *Journal of the Operational Research Society*, *68*(5), 556-565. DOI: https://doi.org/10.1057/s41274-016-0130-2.

Hu, Y., Ma, X., Li, W., Wu, W., & Tu, D. (2020). Forecasting manufacturing industrial natural gas consumption of China using a novel time-delayed fractional grey model with multiple fractional order. *Computational and Applied Mathematics*, *39*(4), 1-30. DOI: https://doi.org/10.1007/s40314-020-01315-3.

Intharathirat, R., Abdul Salam, P., Kumar, S., & Untong, A. (2015). Forecasting of municipal solid waste quantity in a developing country using multivariate grey models. *Waste Management*, *39*, 3-14. DOI: https://doi.org/10.1016/j.wasman.2015.01.026.

Javed, S.A., Zhu, B., & Liu, S. (2020). Forecast of biofuel production and consumption in top CO2 emitting countries using a novel grey model. *Journal of Cleaner Production*, *276*, 123997. DOI: https://doi.org/10.1016/j.jclepro.2020.123997.

Kiran, M., Shanmugam, P.V., Mishra, A., Mehendale, A., & Nadheera Sherin, H.R. (2021). A multivariate discrete grey model for estimating the waste from mobile phones, televisions, and personal computers in India. *Journal of Cleaner Production*, *293*, 126185. DOI: https://doi.org/10.1016/j.jclepro.2021.126185.

Kumar, A., Pant, S., & Ram, M. (2018). Complex system reliability analysis and optimization. In *Advanced Mathematical Techniques in Science and Engineering,* pp. 185-199, River Publisher, Denmark.

Kumar, A., Pant, S., & Ram, M. (2019a). Gray wolf optimizer approach to the reliability-cost optimization of residual heat removal system of a nuclear power plant safety system. *Quality and Reliability Engineering International*, *35*(7), 2228-2239. DOI: https://doi.org/10.1002/qre.2499.

Kumar, A., Pant, S., Ram, M., & Chaube, S. (2019b). Multi-objective grey wolf optimizer approach to the reliability-cost optimization of life support system in space capsule. *International Journal of Systems Assurance Engineering and Management*, *10*(2), 276-284. DOI: https://doi.org/10.1007/s13198-019-00781-1.

Kumar, A., Pant, S., & Singh, S.B. (2017). Reliability optimization of complex systems using cuckoo search algorithm. In *Mathematical Concepts and Applications in Mechanical Engineering and Mechatronics*, IGI Global, USA, pp. 94-110. DOI: https://doi.org/10.4018/978-1-5225-1639-2.ch005.

Lee, Y.S., & Tong, L.I. (2011). Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Conversion and Management*, *52*(1), 147-152. DOI: https://doi.org/10.1016/j.enconman.2010.06.053.

Li, K., & Zhang, T. (2018). Forecasting electricity consumption using an improved grey prediction model. *Information*, *9*(8), 204. DOI: https://doi.org/10.3390/info9080204.

Li, K., & Zhang, T. (2021). A novel grey forecasting model and its application in forecasting the energy consumption in Shanghai. *Energy Systems*, *12*(3), 357-372. DOI: https://doi.org/10.1007/s12667-019-00344-0.

Lin, J., Magnago, F., & Alemany, J.M. (2018). Optimization methods applied to power systems: current practices and challenges. In *Classical and Recent Aspects of Power System Optimization*, pp. 1-18, Academic Press, USA. DOI: https://doi.org/10.1016/B978-0-12-812441-3.00001-X.

Liu, X., & Xie, N. (2019). A nonlinear grey forecasting model with double shape parameters and its application. *Applied Mathematics and Computation*, *360*, 203-212. DOI: https://doi.org/10.1016/j.amc.2019.05.012.

Mirjalili, S., Mirjalili, S.M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, *69*, 46-61. DOI: https://doi.org/10.1016/j.advengsoft.2013.12.007.

Negi, G., Kumar, A., Pant, S., & Ram, M. (2021). GWO: a review and applications. *International Journal of Systems Assurance Engineering and Management*, *12*(1), 1-8. DOI: https://doi.org/10.1007/s13198-020-00995-8.

Pant, S., Kumar, A., & Ram, M. (2017b). Flower pollination algorithm development: a state of art review. *International Journal of Systems Assurance Engineering and Management*, *8*(2), 1858-1866. DOI: https://doi.org/10.1007/s13198-017-0623-7.

Pant, S., Kumar, A., & Ram, M. (2019). Solution of nonlinear systems of equations via metaheuristics. *International Journal of Mathematical, Engineering and Management Sciences*, *4*(5), 1108-1126. DOI: https://doi.org/10.33889/IJMEMS.2019.4.5-088.

Pant, S., Kumar, A., Singh, S.B., & Ram, M. (2017a). A modified particle swarm optimization algorithm for nonlinear optimization. *Nonlinear Studies*, *24*(1), 127-138.

Qu, Z., Mao, W., Zhang, K., Zhang, W., & Li, Z. (2019). Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. *Renewable Energy*, *133*, 919-929. DOI: https://doi.org/10.1016/j.renene.2018.10.043.

Ren, S.J., Wang, C.P., Xiao, Y., Deng, J., Tian, Y., Song, J.J., Cheng, X.J., & Sun, G.F. (2020). Thermal properties of coal during low temperature oxidation using a grey correlation method. *Fuel*, *260*, 116287. DOI: https://doi.org/10.1016/j.fuel.2019.116287.

Shen, X., Yue, M., Duan, P., Wu, G., & Tan, X. (2019). Application of grey prediction model to the prediction of medical consumables consumption. *Grey Systems: Theory and Application*, *9*(2), 213-223. DOI: https://doi.org/10.1108/gs-11-2018-0059.

Sun, W., & Huang, C. (2020). A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network. *Journal of Cleaner Production*, *243*, 118671. DOI: https://doi.org/10.1016/j.jclepro.2019.118671.

Tang, J., Yuan, F., Shen, X., Wang, Z., Rao, M., He, Y., Sun, Y., Li, X., Zhang, W., Li, Y., Gao, B., Qian, H., Bi, G., Song, S., Yang, J.J., & Wu, H. (2019). Bridging biological and artificial neural networks with emerging neuromorphic devices: fundamentals, progress, and challenges. *Advanced Materials, 31*(49), 1902761. DOI: https://doi.org/10.1002/adma.201902761.

Uniyal, N., Pant, S., & Kumar, A. (2020). An overview of few nature inspired optimization techniques and its reliability applications. *International Journal of Mathematical, Engineering and Management Sciences*, *5*(4), 732-743. DOI: https://doi.org/10.33889/IJMEMS.2020.5.4.058.

Wang, C.H., & Hsu, L.C. (2008). Using genetic algorithms grey theory to forecast high technology industrial output. *Applied Mathematics and Computation*, *195*(1), 256-263. DOI: https://doi.org/10.1016/j.amc.2007.04.080.

Wang, R., Xu, K., Xu, Y., & Wu, Y. (2020). Study on prediction model of hazardous chemical accidents. *Journal of Loss Prevention in the Process Industries*, *66*, 104183. DOI: https://doi.org/10.1016/j.jlp.2020.104183.

Wu, W., Ma, X., Wang, Y., Zhang, Y., & Zeng, B. (2019). Research on a novel fractional GM(α , n) model and its applications. *Grey Systems: Theory and Application*, *9*(3), 356-373. DOI: https://doi.org/10.1108/gs-11-2018-0052.

Xuemei, L., Cao, Y., Wang, J., Dang, Y., & Kedong, Y. (2019). A summary of grey forecasting and relational models and its applications in marine economics and management. *Marine Economics and Management*, *2*(2), 87-113. DOI: https://doi.org/10.1108/maem-04-2019-0002.

Yang, Y., Chen, Y., Shi, J., Liu, M., Li, C., & Li, L. (2016). An improved grey neural network forecasting method based on genetic algorithm for oil consumption of China. *Journal of Renewable and Sustainable Energy*, *8*(2), 024104. DOI: https://doi.org/10.1063/1.4944977.

Zeng, B., Li, H., & Ma, X. (2020). A novel multi-variable grey forecasting model and its application in forecasting the grain production in China. *Computers and Industrial Engineering*, *150*, 106915. DOI: https://doi.org/10.1016/j.cie.2020.106915.

Zhang, Q., Yu, H., Barbiero, M., Wang, B., & Gu, M. (2019). Artificial neural networks enabled by nanophotonics. *Light: Science and Applications*, *8*(1), 1-14. DOI: https://doi.org/10.1038/s41377-019-0151-0.