Pamukkale Univ Muh Bilim Derg, 27(3), 384-392, 2021

Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

Evaluation of different curve fitting models for prediction of municipal solid waste composition

Kentsel katı atık bileşiminin tahmini için farklı eğri uydurma modellerinin değerlendirilmesi

Aysun ÖZKAN[1], Kemal ÖZKAN[2], Şahin IŞIK[3], Mufide BANAR[4]*

[1,4]Department of Environmental Engineering, Faculty of Engineering, Eskişehir Technical University, Eskişehir, Turkey.
aysunozkan@eskisehir.edu.tr, mbanar@eskisehir.edu.tr

[2,3]Department of Computer Engineering, Faculty of Engineering and Architecture, Eskişehir Osmangazi Univ., Eskişehir, Turkey.
kozkan@ogu.edu.tr, sahini@ogu.edu.tr

Received/Geliş Tarihi: 04.07.2019 Revision/Düzeltme Tarihi: 26.03.2020 doi: 10.5505/pajes.2020.08377
Accepted/Kabul Tarihi: 17.04.2020 Research Article/Araştırma Makalesi

Abstract

In this study, the methodology was applied for the prediction of municipal solid waste (MSW) composition in the Eskişehir/Turkey. For this purpose, MSW samples were collected, and the samples were separated by food wastes, paper-cardboard, plastics, glass, metals as manually. In the present study, the concept of 2D curve fitting functions was adopted for the forecasting of MSW composition. To comment on the performance of proposed system, the Root Mean Square Error (RMSE) and Sum of Squared Error (SSE) metrics were chosen as statistical residual evaluation metrics. According to results, it is seen that the polynomial curve fitting model is most suitable. Also, the effect of socio-economic structure on the waste composition by different nonlinear models was observed. With contribution of study, it would be possible to forecast the MSW composition of other cities which have similar factors.

Keywords: Curve fitting, Modeling, Municipal solid waste, Socioeconomic factors, Waste composition.

Öz

Bu çalışmada, Eskişehir/Türkiye'deki kentsel katı atık (MSW) bileşiminin tahmini için bir metodoloji uygulanmıştır. Bu amaçla, MSW örnekleri toplanmış ve numuneler yiyecek atığı, kâğıt-karton, plastik, metal, cam olarak el ile ayrılmıştır. Bu çalışmada, MSW bileşiminin tahmini için 2B eğri uydurma fonksiyonları kullanılmıştır. Önerilen sistemin performansını yorumlamak için, Kök Ortalama Karesel Hata (RMSE) ve Toplam Karesel Hata (SSE) değerleri değerlendirme ölçütleri olarak seçilmiştir. Sonuçlara göre, polinom eğri uydurma modelinin atık bileşiminin tahmini için daha uygun olduğu belirlenmiştir. Ayrıca, sosyoekonomik yapının atık bileşimi üzerindeki farklı doğrusal olmayan modellerle etkisi gözlenmiştir. Çalışmanın katkısıyla, benzer faktörlere sahip olan diğer şehirlerin MSW kompozisyonunu tahmin etmek mümkün olacaktır.

Anahtar kelimeler: Eğri uydurma, Modelleme, Kentsel katı atık, Sosyoekonomik faktörler, Atık bileşimi.

1 Introduction

The municipal solid waste (MSW) is very important since the contained ingredients can directly cause negative human health effects and environmental pollution. It is widely known that MSW can lead to the proliferation of insects, passing of the harmful instances into the plants, foodstuffs, and spread of epidemic diseases and etc. Specially, the decomposition of glass and plastics takes many years and causes pollution of soil permanently. To reduce problems encountered from MSW, a sophisticated solution for municipalities is greatly needed in terms of facilitating the reduction, reuse, recycling, and disposal stages. Therefore, the general trend behind a sufficient and accurate projected sample can be utilized to make decisions about the size and quantity of municipal solid wastes.

The emerging works in MSW monitoring deal with methodologies and techniques in order to determine the location of organic recycling and processing plants. It has been emphasized that there are two major ideas as time series and factor models that can be used for the prediction of the volume and future dimension MSW generation [1]. The key idea under the time series models is considered that the obtained data within a period can be modeled based on the trend or autocorrelation between consecutive data points. In this respect, the time series analysis can be carried out to extract meaningful information by considering the characteristic of the present data. In this study, we will review only the techniques that have been utilized for forecasting of MSW composition of different locations as using the projected data. Typical methods in time series are curve fitting, exponential smoothing and auto correlation (Auto-Regressive Integrated Moving Average (ARIMA)) models. The first study on ARIMA model was proposed by Box and Jenkins [2]. Previously, the limitations of S-curves fitting methodology were investigated to analyze for long-term forecasting of MSW composition in the UK [3]. Mwenda el al. [4] have been utilized the ARMA/ARIMA and exponential smoothing models to predict the quantity of solid waste generation in Arusha City-Tanzania. Also, Sodanil and Chatthong [5] have been operated Artificial Neural Network (ANN) models based on time series for prediction solid waste generation for Bangkok city of Thailand. Additionally, two forecasting techniques as seasonal Auto Regressive and Moving Average (sARIMA) model [2] and a discrete dynamical system for analysis of non-linear systems in order to predict MSW generation rates, based on the MSW time series of three cities in Spain and Greece [6].

*Corresponding author/Yazışılan Yazar

Another technique considers socio-economic factors (age, educational level, and income level) as basis for forecasting MSW production rate. The aim of factor-based methods is unveiling the relationship between factors and generated waste composition. In a work, a multivariate regression model is developed to disclose the meaningful relationship between the demographic parameters and the amount of collected waste paper [7]. For this purpose, a hypothesis is put forward that the waste paper potential and factors influence the size of collected paper and density of collection sites. Again, Ying et al. [8] tried to model the projected data with coefficients of Gray forecast and multiple linear regression model with respect to Minimal Sum of Square Error (MSSE). Then, the rate of municipal solid waste from 2010 to 2014 was estimated based on the statistics data of Beijing satellite towns in 1992-2005. The grey fuzzy dynamic modeling is carried out for the prediction of MSW generation a case study in the city of Beijing in China [9]. The authors emphasized that the new forecasting technique should be preferred instead of the conventional grey dynamic model, least squares regression approach and the fuzzy goal regression model in terms getting the best accuracy rates. Similarly, Dai et al. [10] utilized the two-stage support vector regression optimization model (TSOM) for the monitoring of MSW management in the urban districts of Beijing. In a different study, optimized multivariate grey model was applied to forecast MSW collected in Thailand with prediction intervals in long term period [11]. The potential impact of coupling the support-vector-regression (SVR) model with an interval parameter mixed-integer linear programming (IMILP) was investigated to optimize the MWM planning. To evaluate the accuracy of system, three types of performance evaluation criteria such as prediction accuracy (PA), fitting accuracy (FA) and overall accuracy (OA determined by employing the four kernel functions including linear kernel, polynomial kernel, radial basis function, and multilayer perception kernel in TSOM). Then, the SVR model with the best kernel function was chosen to forecast the waste generation rate in Beijing.

The composition of municipal solid waste (MSW) is a result of regional and cultural aspects as well as social behavior, and it is strongly influenced by economic factors. The changes in MSW composition may strongly influence the quality of the waste, which affects emissions from landfills, the quality of incineration residues and other parameters of waste management systems [12]. Moreover, Denafas et al. [12] analyzed the effects of seasonal variation in MSW composition by relying on the time series forecasting models including non-parametric seasonal exponential smoothing, Winters additive, and Winters multiplicative methods on collected data. Furthermore, the potential impact of ANN was examined to provide a convincing accuracy rate in terms of MSW monitoring. Specifically, Zade and Noori [13] addressed the ability of the ANN for forecasting of MSW generation in Mashhad. Kumar et al. [14] have been estimated the amount of MSW in the Eluru/India with the ANN model. Also, ANN models have been used for MSW prediction in Gujarat (India) by Patel and Meka [15].

In fact, the prediction of the MSW generation task is similar to determine the severity of a possible incoming earthquake or the weather conditions of a next week, since for both tasks a model is constructed based on the data collected within a period in the past. Usually, using an accurate mathematical modeling technique can promise satisfactory results for expectations. Besides, it is accepted that some favorable and unfavorable aspects of methods are required to be accounted for different conditions. For instance, when the dimension of training data is increased in case of ANN based modeling, some problems also arise in the sense of determination of network architecture, local minimum and parameter selection [11],[16]. Moreover, 1D curve functions have been utilized to make predictions on MSW generation in our previous study [17]. Similarly, one should make a tradeoff between the accuracy and computation time if the time series based approaches are utilized. Moreover, there is no accepted unique method that produces the best accuracy rates in terms of forecasting MSW generation for different data types of different municipalities.

Eskişehir which is studies city in this paper has a land area of 13.925 km^2 and is located in the northwest of the Central Anatolia region in Turkey. The city is also known as a university town. According to the report of 2014 on the census, the population of the city is around 810,000 and the population density is about 58/km^2. There are two municipalities in Eskişehir as Odunpazari and Tepebasi. Along with student population growth, the quantity of consumption has also increased. The statistical record indicates that the average daily MWS production rate is reported as 750 tons/day. At some certain times of the day, two private companies are organized for collecting the MSW in plastic bags. The collected MSW is transferred into a landfill.

In this study, it was realized a feasibility work that serves as extracting the meaningful information from historical statistical data of MSW generation for two municipalities including Odunpazari and Tepebasi of Eskişehir city in Turkey. With this aim, it was conducted some experiments on a sample of projected MSW generation by investigating the limitation of two dimensional (2D) curve fitting methods, which are Power, Exponential and Polynomial. The reason to choose these curve types is explained as the aforementioned curve types are faster than other ones in terms computation time and constructing the model from large scale data. According to the literature of predictions for MSW generation rate, it is observed that while the one-dimensional (1D) curve fitting strategies have been applied for forecasting of MSW generation ([18] and [19]) , but the two-dimensional (2D) based methods have not attempted yet for this purpose. Technically, the 1D curve fitting model implies that there is only one independent variable. Also, in the 2D case, there are two different independent variables. Moreover, there is no work that proposed to make prediction for MSW generation in Eskişehir. The objectives of this study are (i) to analyze the relationship between population and income level with food wastes, paper-cardboard, plastics, glass, metals, and others composition types and (ii) to investigate the ability to make predictions for different municipality and countries that share the same characteristics in socio-economic factors.

2 Material and method

The motivation behind the curve fitting is constructing a mathematical function that captures the trend in data and the relationship between the factors (independent) and outcomes (dependent). For this purpose, the parameters of the fitted function are adjusted as yielding the 'best-fit' of the model. In determining the characteristic of data and obtaining the best-fitted model, linear and nonlinear curve fitting functions can be preferred. The designation of subject is ground on as interpolation or smoothing. In this study, we have extended the

curve fitting strategy for interpolation purposes in order to predict the quantity of MSW generation.

In the power-based strategy, the mathematical formula given in Equation (1) is utilized.

$$p(x,y) = a_0 + a_1 x^{a_2} + a_3 y^{a_4} + a_5 x^{a_6} y^{a_1} \quad (1)$$

In the case of exponential based curve fitting, the model denoted with Equation (2) is aimed to extract by using the factors and dependent variables. In given equation, the a_0, a_1 and a_2 refer to coefficients need to be determined, whereas μ_x and μ_y denote the mean of variables corresponding to the x and y planes.

$$p(x,y) = a_0 e^{-\left[\frac{(\bar{x}-\mu_x)^2}{2a_1^2} + \frac{(\bar{y}-\mu_y)^2}{2a_2^2}\right]} \quad (2)$$

Where the parameters are given as $\bar{x} = xcos(\theta) - ysin(\theta), \bar{y} = xsin(\theta) - ycos(\theta), \quad \mu_x = a_3 cos(\theta) - a_4 sin(\theta)$ and $\mu_y = a_3 sin(\theta) - a_4 cos(\theta)$.

If our aim is modeling the data by a second-order polynomial function (Equation 3);

$$y = a_0 + a_1 x_1 + a_2 x_2 + e \quad (3)$$

where e is the expected residual value, Equation 4 can be used [20].

$$e_{sum} = \sum_{i=1}^{m} (y_i - a_0 - a_1 x_i - a_2 x_i^2)^2 \quad (4)$$

Then, the derivative of both sides of the equation should be derived as in Equation (5);

$$\begin{aligned} \frac{\partial e_{sum}}{\partial a_0} &= -2\sum_{i=1}^{m} (y_i - a_0 - a_1 x_i - a_2 x_i^2)^2 \\ \frac{\partial e_{sum}}{\partial a_1} &= -2\sum_{i=1}^{m} x_i (y_i - a_0 - a_1 x_i - a_2 x_i^2) \\ \frac{\partial e_{sum}}{\partial a_2} &= -2\sum_{i=1}^{m} x_i^2 (y_i - a_0 - a_1 x_i - a_2 x_i^2) \end{aligned} \quad (5)$$

So, the generated functions can be rearranged (Equation 6):

$$\begin{bmatrix} n & \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i^3 \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^3 & \sum_{i=1}^{m} x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \sum_{i=1}^{m} x_i^2 y_i \end{bmatrix} \quad (6)$$

By extending the concept of Equation 3, one can easily develop m th-order polynomial curve fitting as shown in Equation (7):

$$\begin{bmatrix} n & \sum_{i=1}^{m} x_i & \dots & \sum_{i=1}^{m} x_i^2 \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \cdot & \sum_{i=1}^{m} x_i^3 \\ \cdot & \cdot & \cdot & \cdot \\ \sum_{i=1}^{m} x_i^m & \sum_{i=1}^{m} x_i^{m+1} & \dots & \sum_{i=1}^{m} x_i^{m+m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ a_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \cdot \\ \sum_{i=1}^{m} x_i^m y_i \end{bmatrix} \quad (7)$$

In the present study, the concept of 2D curve fitting functions is adopted for forecasting of MSW generation. In other saying, the relation between dependent variables and two factors (income and population) is investigated by using the 2D curve fitting functions, respectively. In the sense of accurate prediction, one can emphasize that the contribution of 2D based curve fitting methodology is more convenient than the other one when the minimum error and stability are considered as the principal constraints. This can be attributed to the impact of 2D based curve fitting methodology as analyzing the rising and declining property of data from two aspects, by generating a surface based on the two factors. The aforementioned curve fitting methods are power, exponential and polynomial.

Let assume that we have given two factors as x and y to build a model on processed data. In this study, the population (x) and income level (y) are considered as two factors that influence the variation on dependent variables, which are food wastes, paper-cardboard, plastics, glass, metals, and others. The given two variables reveal the effect of socio-economic factors on MSW prediction. For polynomial case, the model is constructed with Equation (8).

$$p(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{ij} x^i y^j \quad (8)$$

Where n and m are the degree of x and y factors, respectively. In this study, n and m values is equal to 3.

As touched above, the parameters of the model can be easily derived by carrying out the derivative procedure in terms of obtaining the 'best-fitted model' for 1D curve fitting functions. However, for the 2D based curve fitting concept, the optimization problem can be solved with a nonlinear optimization technique. To solve the optimization for nonlinear functions, the constrained nonlinear optimization is employed. The constrained nonlinear optimization for 2D functions has the form as indicated in Equation (9).

$$\min_{x,y\in^{\sim}n} \sum_{i=1}^{6} (f_i(x,y) - z_i)^2$$
$$subject\ to \quad \sum_{i=1}^{6} f_i(x,y) = 100 \tag{9}$$

Where, x and *y* represent the population and income level, and the value of *i* varies between 1 to 6. Also, the z_i denotes percentage of variation for dependent variables including food wastes, paper-cardboard, plastics, glass, metals, and others. The motivation under using the nonlinear optimization is explained with extracting the coefficients of 2D based curve fitted functions by taking the residual minimization constraint. Since the sum of the probabilities of waste generation is given as percentage, the total sum of all waste products should also yield the 100. To solve the constrained nonlinear optimization, the lsqcurvefit function that implemented in Matlab optimization toolbox, is carried out. The referred function is utilized to solve nonlinear curve-fitting (data-fitting) problems in least-squares sense.

To comment on the performance of the proposed system, the Root Mean Square Error (RMSE) and Sum of Squared Error (SSE) metrics are chosen as statistical residual evaluation metrics [21]. As a numerical prediction, the RMSE is the modified version of mean square error and presents the difference between the estimated and expected value. The general formula used to compute the RMSE value as shown in Equation (10).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{observed,i} - X_{predicted,i})^2}{n}} \tag{10}$$

The SSE is computed with formula shown in Equation (11).

$$SSE = \sqrt{\sum_{i=1}^{n}(X_{observed,i} - X_{predicted,i})^2} \tag{11}$$

In this study, the accuracy of system is calculated based on the SSE and RMSE values. Thereby, the real data is put forward into the constructed model and the predictions for compositions are obtained. Then, the residual values with SSE and RMSE are obtained based on the predicted and observed results. The smaller the SSE and the RMSE values, the more accurate the prediction of unknown values.

In this study, 293 MSW samples were collected from different regions at different seasons. The amount of each component was determined gravimetrically after manual sorting [22],[23]. The population data were obtained from the Turkish Statistical Institute. The income level values were entered according to land market value in the sub municipalities' databases.

To investigate the relationship between socio-economic factors and MSW compositions, the different types of curves and loss metrics are carried out in connection with prediction. After providing a limited sample for each factor and composition, the variation in income and population versus food wastes, paper-cardboard, plastics, glass, metals are examined based upon the constructed mathematical models. With this aim, models are determined based on the obtained coefficients of fitted curves, which are power, exponential and polynomial curve fitting methods. As an optimization problem, the minimum RMSE and SSE error values are considered as high accuracy rate for forecasting of MSW generation. For the sake of accurate prediction, only 2D curve fitting concept is utilized in the experimental stage. Conceptually, the residual between the predicted 2D plane and observed 2D plane is accounted to fulfill the forecasting of MSW generation with plausible accuracy.

3 Results and Discussion

In case of assessing the performance of system, the simple leave-one-out cross-validation methodology is applied in experiments from 1 to 3. As a degenerate version of k-fold cross-validation approach, the value of k is selected as 1 in case of cross-validation. Hence, among the n districts of aforementioned municipalities, the samples of n-1 are processed for model construction and remaining one for prediction procedure purposes, until all districts are completed. Moreover, the generation of MSW compositions for Tepebasi (sub municipality A) is predicted based upon the model obtained from Odunpazari (sub municipality B) data, and vice versa. Once all of the cross-validation stages completed, an overall residual error for each MSW composition is obtained. The reason of realize the experiments 1 and 2 is cross-validated the performance of proposed system for predicting MSW composition of one municipality based on the data of other municipality. Since the real value of MSW generation for both municipalities are available, we can easily compute the performance of system with respect to SSE and RMSE values.

3.1 Experiment 1: Prediction of MSW generation for sub municipality B with sub municipality a data

In this experiment, the amount of MSW generation for sub municipality B was aimed to forecast based on the mathematical models constructed from projected samples of sub municipality A. Thereby, 2D type of Power, Exponential and Polynomial curve fitting functions are utilized to constitute three different models on the sub municipality B data. By experimenting with the models and error metrics (SSE and RMSE), the predictions have been done and the system was evaluated in terms of overall accuracy.

The obtained error rates and formulations for MSW generation in sub municipality B are given in Table 1. By taking the average of the given of SSE values in Table 1, one can observe that the constructed models present 6.84, 18.44, 10.78, 8.80, 13.44 and 16.93 residual values for versus paper-cardboard, metal, glass, plastic, food waste and others, respectively. Also, in case of RMSE metric, the obtained average residual values reach the minimum values as 1.11, 3.01, 1.77, 1.45, 2.22 and 2.76 for paper-cardboard, metal, glass, plastic, food waste and others, respectively. When both values are evaluated, it is seen that the paper-cardboard component has the lowest error value. Moreover, it can be emphasized that different curve functions are suitable to make predictions for each composition when the SSE and RMSE error metrics ground on. Specifically, Exponential, Power, Power, Exponential, Exponential, Exponential and Power are more convenient to forecast MSW generation for food wastes, paper-cardboard, plastics, glass, metals, and others, respectively, if the SSE is considered as reference. However, in case of lowest RMSE, the best-fitted curve functions are noted as Polynomial, Power, Power, Polynomial, Exponential and Power for paper-cardboard, metal, glass, plastic, food waste and others, respectively.

Table 1. Predictions for sub municipality B based on the sub municipality A data.

	Waste type	Functions	SSE	RMSE
Power	Paper-cardboard	$2.0610x_{income}^{-0.3858} + 9.1377x_{population}^{0.0568} + 14.3527x_{income}^{0.1344}x_{population}^{0.0744} - 14.1310$	12.17	2.30
	Metal	$7.1443x_{income}^{0.2289} + 6.0369x_{population}^{0.0721} - 12.0389x_{income}^{1.4219}x_{population}^{0.4326} - 7.7760$	5.67	1.07
	Glass	$3.5688x_{income}^{-0.0545} + 5.9787x_{population}^{-0.0465} + 19.8102x_{income}^{2.4718}x_{population}^{2.5146} - 7.8847$	7.21	1.36
	Plastic	$1.3400x_{income}^{12.2168} + 16.1529x_{population}^{-0.0695} + 15.9471x_{income}^{0.6080}x_{population}^{6.3378} - 9.1839$	9.62	1.82
	Food waste	$156.5496x_{income}^{-0.0052} - 7.0024x_{population}^{1.6351} + 0.0011x_{income}^{-1.1443}x_{population}^{-4.2121} - 91.8003$	15.81	2.99
	Others	$12.2014x_{income}^{0.1542} + 37.2973x_{population}^{0.0186} + 36.5557x_{income}^{0.4752}x_{population}^{-0.0757} + 32.3670$	14.14	2.67
	Average		10.77	2.03
Exponential	Paper-cardboard	$a_0 = 271.6647, a_1 = -64.7208, a_2 = 25.9419, a_3 = 6.5275, a_4 = 1.0000, \theta = 0.0846$	12.11	2.29
	Metal	$a_0 = 203.5130, a_1 = -6.6165, a_2 = 67.4765, a_3 = 10.6972, a_4 = 4.7197, \theta = 0.1910$	5.82	1.10
	Glass	$a_0 = 203.5130, a_1 = -6.6165, a_2 = 67.4765, a_3 = 10.6972, a_4 = 4.7197, \theta = 0.1910$	8.35	1.58
	Plastic	$a_0 = 1804.6, a_1 = 3.4321, a_2 = 163.8965, a_3 = 6.8159, a_4 = 2.1201, \theta = 0.3937$	9.43	1.78
	Food waste	$a_0 = 3926.8, a_1 = -841.3752, a_2 = 1.6228, a_3 = -1045.2, a_4 = 471.3370, \theta = 0.6781$	15.63	2.95
	Others	$a_0 = 49.3872, a_1 = 193.9565, a_2 = 198.4946, a_3 = -193.3062, a_4 = 0.4057, \theta = 0.7854$	14.65	2.77
	Average		10.99	2.08
Polynomial	Paper-cardboard	$-3.1198x_{income}^2 + 2.7903x_{income} + 5.5966 + 32.8721x_{population}^2 - 20.7708x_{income}x_{population} + 14.0527x_{population} - 51.4347x_{income}^2x_{population}^2 + 47.1411x_{income}x_{population}^2 - 5.2481x_{population}^2$	16.03	2.07
	Metal	$-1.9136x_{income}^2 - 1.2016x_{income} + 2.2648 - 6.7318x_{population}^2 + 4.1461x_{income}x_{population} - 1.2127x_{population} + 11.8687x_{income}^2x_{population}^2 - 11.9685x_{income}x_{population}^2 + 2.5561x_{population}^2$	9.02	1.16
	Glass	$-0.6967x_{income}^2 + 1.4068x_{income} + 1.1887 + 8.5279x_{population}^2 - 2.6598x_{income}x_{population} + 1.2809x_{population} - 10.3535x_{income}^2x_{population}^2 + 6.2096x_{income}x_{population}^2 + 0.2209x_{population}^2$	10.84	1.40
	Plastic	$-2.2162x_{income}^2 + 3.7103x_{income} + 3.8389 + 12.7237x_{population}^2 - 3.9512x_{income}x_{population} - 1.4814x_{population} - 1.2552x_{income}^2x_{population}^2 - 9.6663x_{income}x_{population}^2 + 8.9875x_{population}^2$	13.28	1.71
	Food waste	$5.4586x_{income}^2 + 0.5325x_{income} + 68.8942 + 51.8570x_{population}^2 + 31.0180x_{income}x_{population} - 4.5711x_{population} + 44.7518x_{income}^2x_{population}^2 - 34.6319x_{income}x_{population}^2 - 7.6441x_{population}^2$	23.89	3.09
	Others	$-4.7868x_{income}^2 - 0.3485x_{income} + 16.4734 - 2.3476x_{population}^2 - 10.4435x_{income}x_{population} - 5.2611x_{population} + 4.2848x_{income}^2x_{population}^2 + 1.2804x_{income}x_{population}^2 - 2.7320x_{population}^2$	22.00	2.84
	Average		15.84	2.05

Upon inspecting the overall results, it can be observed that the performance of Power is favorable than the other ones. One can find that the proposed system presents 2.03 RMSE residual as the overall error rate. This fact validates that proposed models can be used accurately for prediction MSW generation. The small SSE and RMSE values imply that the proposed system is useful to make prediction on MSW generation for any municipality in Eskişehir.

3.2 Experiment 2: Prediction of MSW generation for sub municipality B with sub municipality B data

To evaluate the ability of proposed prediction system, the waste composition rate of sub municipality A is estimated by utilizing the models computed from sub municipality B data. Similarly, the aforementioned 2D curve types including Power, Exponential and Polynomial are executed to build models in order to obtain some estimation for MSW generation in sub municipality A. The performance of the system is analyzed bases upon the SSE and RMSE error metrics. The residual values and formulations corresponding to sub municipality A are demonstrated in Table 2. The average error values associated with SSE are reported to 10.79, 6.99, 7.99, 9.26, 17.22 and 17.05 while the average RMSE outputs for prediction are reported to 1.91, 1.24, 1.41, 1.64, 3.05, 3.02 and 2.04, for the prediction of paper-cardboard, metals, glass, plastics, food wastes and others compositions, respectively.

Table 2. Predictions for sub municipality A based on the sub municipality B data.

	Waste type	Functions	SSE	RMSE
Power	Paper-cardboard	$-43.3522x_{income}^{20.0835} + 66.8376x_{population}^{0.0949} + 10.1771x_{income}^{7.7103}x_{population}^{-0.6290} - 47.7461$	11.39	2.01
	Metal	$22.6090x_{income}^{0.1789} + 31.0098x_{population}^{-1.7439} - 11.5861x_{income}^{0.6705}x_{population}^{0.0110} - 11.1613$	6.82	1.21
	Glass	$11.2887x_{income}^{0.0713} + 9.3134x_{population}^{0.1049} + 9.2882x_{income}^{0.0768}x_{population}^{0.00193} - 13.8346$	8.10	1.43
	Plastic	$7.3142x_{income}^{0.6960} + 8.2364x_{population}^{3.3410} - 58.1417x_{income}^{9.8529}x_{population}^{1.1943} + 2.1180$	9.23	1.63
	Food waste	$16.9448x_{income}^{-0.0910} + 50.9708x_{population}^{-0.3263} - 12.5065x_{income}^{-0.0421}x_{population}^{-0.6378} + 0.7269$	17.63	3.12
	Others	$-78.9470x_{income}^{-0.8596} + 16.7113x_{population}^{0.1233} + 13.5468x_{income}^{0.1708}x_{population}^{0.9577} + 71.2642$	17.24	3.05
	Average		11.74	2.08
Exponential	Paper-cardboard	$a_0 = 11.9404, a_1 = 0.6323, a_2 = 3.7318, a_3 = 1.4783, a_4 = 1.8110, \theta = 0.1454$	10.44	1.85
	Metal	$a_0 = 3.5848, a_1 = -2.3488, a_2 = 2.4660, a_3 = 1.7464, a_4 = 3.8130, \theta = 0.3268$	7.12	1.26
	Glass	$a_0 = 5.2215, a_1 = 2.8581, a_2 = 2.2755, a_3 = 1.8886, a_4 = 3.6301, \theta = -0.2710$	7.98	1.41
	Plastic	$a_0 = 8.2816, a_1 = -2.4781, a_2 = 3.9868, a_3 = 0.9544, a_4 = 6.2709, \theta = 0.5093$	9.77	1.73
	Food waste	$a_0 = 161.0675, a_1 = -5217.9, a_2 = 4192.1, a_3 = 1092.5, a_4 = 6.6314, \theta = 0.2068$	16.92	2.99
	Others	$a_0 = 18.5377, a_1 = -4.0483, a_2 = 5.6646, a_3 = -5.8217, a_4 = 8.8031, \theta = 0.7560$	16.83	2.98
	Average		11.51	2.04
Polynomial	Paper-cardboard	$0.0869x_{income}^2 - 1.1383x_{income} + 7.4797 + 1.1086x_{population}^2 - 0.2785x_{income}x_{population} + 5.0933x_{population} - 5.4369x_{income}^2x_{population}^2 + 1.7445x_{income}x_{population}^2 - 1.0876x_{population}^2$	10.55	1.87
	Metal	$-0.9572x_{income}^2 - 0.6267x_{income} + 1.6580 - 5.1448x_{population}^2 + 2.5605x_{income}x_{population} - 0.3468x_{population} + 10.4910x_{income}^2x_{population}^2 - 6.6061x_{income}x_{population}^2 + 0.9486x_{population}^2$	7.03	1.24
	Glass	$0.4271x_{income}^2 - 1.4925x_{income} + 3.2939 - 0.2813x_{population}^2 + 3789x_{income}x_{population} - 1.4572x_{population} + 4.9299x_{income}^2x_{population}^2 + 0.3702x_{income}x_{population}^2 - 0.5985x_{population}^2$	7.88	1.39
	Plastic	$3.3700x_{income}^2 - 2.5399x_{income} + 4.2115 - 18.8385x_{population}^2 + 16.6522x_{income}x_{population} + 0.8925x_{population} + 33.3230x_{income}^2x_{population}^2 - 21.5213x_{income}x_{population}^2 + 7.7760x_{population}^2$	8.78	1.55
	Food waste	$-3.2050x_{income}^2 + 5.9104x_{income} + 65.3479 - 6.6904x_{population}^2 + 13.9028x_{income}x_{population} - 3.0416x_{population} - 91.4795x_{income}^2x_{population}^2 + 23.7322x_{income}x_{population}^2 - 8.3663x_{population}^2$	17.12	3.03
	Others	$-5.6125x_{income}^2 + 6.3669x_{income} + 16.2643 + 28.9078x_{population}^2 - 34.5171x_{income}x_{population} + 4.5028x_{population} + 47.1655x_{income}^2x_{population}^2 + 0.8307x_{income}x_{population}^2 - 3.5068x_{population}^2$	17.08	3.02
	Average		11.41	2.02

Again, it can be noted that different curve types present different performances for MSW prediction. Especially, we can see that Exponential, Power, Polynomial, Polynomial, Exponential and Exponential curve types are more convenient for prediction of paper-cardboard, metals, glass, plastics, food wastes and others compositions, respectively, when the SSE residual values are considered. However, in case RMSE metrics, the Exponential, Power, Polynomial, Polynomial, Exponential, and Exponential curve types should be utilized in terms of forecasting the paper-cardboard, metals, glass, plastics, food wastes and others compositions, respectively. Moreover, we can say that the Polynomial curve types are more appropriate for all types of MSW composition in Eskişehir if the overall error metrics are considered. Likewise to sub municipality B, the overall average error values are reported as 11.41 and 2.02 in case of SSE and RMSE metrics. Based on the obtained good accuracy results, one can note that it possible to predict MSW rate of a municipality in Eskişehir.

3.3 Experiment 3: Performance evaluation on all data

In this experiment, the MSW quantity of both sub municipality A and sub municipality B are combined to make more accurate predictions for MSW generation in Eskişehir. For this purpose, the prediction of food wastes, paper-cardboard, plastics, glass, metals, and others. MSW composition associated with each district are determined in terms of SSE and RMSE error values. In general, the utilized accuracy evaluation system is named as leave-one-out in literature that is a large proportion of data is experimented to construct the model whereas remain proportion is utilized for testing (prediction) purpose. Finally, the SSE and RMSE values are obtained.

According to the results in Table 3 obtained from leave-one-out system, it can be emphasized that curve type should be utilized in the prediction of food wastes, paper-cardboard, plastics, glass, metals and others composition, respectively, if the SSE values are ground on. Again, for RMSE values, it is obviously seen that Power, Polynomial, Polynomial, Polynomial, Exponential, Exponential and Power are useful for MSW composition prediction particularly for paper-cardboard, metals, glass, plastics, food wastes and others compositions, respectively. For overall evaluation, we have found that Polynomial curve fitting is best for the prediction of waste composition types, especially packaging wastes. Moreover, the overall average values are reported as 15.44 and 2.04 for SSE and RMSE error metrics, respectively. By considering the obtained small SSE and RMSE values, it can be emphasized that the proposed system is useful to predict the MSW generation in other cities/countries that share similar characteristic with Eskişehir when the training data related to past season are available.

3.4 Experiment 3: Performance evaluation on all data

Multiple linear regression (MLR) includes many advantages to find an accurate model that reveals the relation between generated MSW rate and social-economic factors. To compare the limitations of some similar studies, we have enhanced the performance evaluation and comparison stage by justifying their usefulness. Since the existing methods are task-oriented and considered different datasets, it would be unfair to compare performances.

One can say that the ANN has indisputable advantages to find a best model capturing the characteristic of data. In a given study [24], the valuable RMSE score, which accounts for 467 was obtained in case of in predicting the municipal solid waste generation. Again, the double exponential smoothing models [4] were utilized to find out the best time series model for forecasting amount of solid waste generation for a following year. The obtained RMSE is noted as 348.60. In a similar study [12], the seasonal variation of municipal solid waste generation and composition for Kaunas, Lithuania, was predicted with a 0.371 of RMSE value. Recently, an algorithm with a similar concept behind our approach is proposed by Kannangara et al. [25]. In related study, the prediction of regional municipal solid waste generation was investigated by considering Socio-economic variables including population, median personal income, employment rate and etc. After setting parameters of ANN, the model gives a satisfactory prediction accuracy with 28.996 of MSE value when testing system over income and population as influenced factors. Moreover, the multiple linear regression model [26] was used on the basis of social and demographic explanatory variables such as daily per capita income and population. In related study, the developed multiple linear regression model shows the valuable prediction result, recorded as 0.19 of RMSE value. One can note that the main idea is representing the characteristic of data with a robust model. In order to analyze the merits and limitations of some curve fitting models on MSW prediction, we have also carried out various experiments. After experimental stage, it turns out that the power model gives superior results than other model. At overall, the obtained RMSE values implies that the curve fitting based models give accurate predictions to reveal the relationship between MSW generation and independent factors like population and income.

4 Conclusion

In this paper, a new perspective is presented for the utilization of some curve fitting functions for prediction of MSW composition. By considering the obtained error rates, we consider it necessary to recommend some suggestions in the sense of taking some precautions in terms of planning a solid waste management system. By considering the experiment 1 to 3, we can comment that it is useful to combine population and income factors with a purpose to predict the amount of MSW generation in Eskişehir, instead of using one factor. Particularly, predicting the MSW generation of metal gives best accuracy rates when all experiments are considered. While making predictions for paper-cardboard, glass, plastic compositions gives good accuracy in terms of SSE and RMSE metrics, however, it is a big challenge to predict the food waste and generation of others compositions according to SSE and RMSE. One can list such possible reasons as high variation within variables and low connection between variables and outcomes. Given the small RMSE and SSE values, it is clear that the curve fitting models can be used for waste composition estimation. When all the results are evaluated together, it is seen that the polynomial curve fitting model is most suitable. With contribution of study, it would be possible to forecast the MSW composition of other cities which have similar factors (income level and population) in common. Therefore, one can emphasize that the curve fitting based data-driven methodology can be used with a purpose of generating predictions for other municipalities. In order to enhance the prediction capability of a model, a proposed model must be retrained to receive valuable forecasting results in case of simulations. For an unseen data, the model have to be generalized in order capture the trend behind newly variables.

Table 3. Overall prediction by conducting leave one out model on all data.

	Waste type	Functions	SSE	RMSE
Power	Paper-cardboard	$1.8826x_{income}^{1.8859} + 9.2304x_{population}^{0.0926} + 6.9212x_{income}^{0.0508}x_{population}^{0.1077} - 3.1313$	15.90	2.05
	Metal	$7.3049x_{income}^{-0.1065} - 1.83918x_{population}^{-0.1134} + 8.6879x_{income}^{-0.0825}x_{population}^{0.0204} - 12.7587$	12.54	1.62
	Glass	$10.8099x_{income}^{0.1667} + 1.3955x_{population}^{-0.1606} - 3.9373x_{income}^{0.4762}x_{population}^{-0.1533} - 4.6410$	11.61	1.50
	Plastic	$4.9295x_{income}^{0.0793} + 5.9723x_{population}^{5.1522} + 7.4617x_{income}^{0.0840}x_{population}^{0.2107} - 3.4608$	13.51	1.75
	Food waste	$48.5099x_{income}^{0.0486} - 21.9386x_{population}^{-0.4984} - 31.4466x_{income}^{-0.0449}x_{population}^{-0.4315} + 0.005$	24.17	3.12
	Others	$-3.3167x_{income}^{-0.1567} - 7.7968x_{population}^{2.5409} + 10.2458x_{income}^{0.5167}x_{population}^{0.3578} + 7.1547$	20.17	2.60
	Average		16.32	2.11
Exponential	Paper-cardboard	$a_0 = -8.0342, a_1 = 1.3904, a_2 = 4.8707, a_3 = 1.4478, a_4 = 3.0503, \theta = -0.7577$	16.01	2.07
	Metal	$a_0 = 12.6625, a_1 = -0.6450, a_2 = -1.4157, a_3 = 2.6330, a_4 = 1.4639, \theta = 0.3890$	10.01	1.29
	Glass	$a_0 = 6.0364, a_1 = 1.7130, a_2 = 2.5366, a_3 = 0.4925, a_4 = 1.7336, \theta = -0.4427$	11.48	1.48
	Plastic	$a_0 = 310.8355, a_1 = -147.7572, a_2 = 5.0740, a_3 = -539.7156, a_4 = 204.5292, \theta = 0.2676$	14.32	1.85
	Food waste	$a_0 = 61.9604, a_1 = -3.4387, a_2 = 107.0465, a_3 = -29.6891, a_4 = 91.5421, \theta = 0.7109$	23.32	3.01
	Others	$a_0 = 17.0357, a_1 = -1.3116, a_2 = 13.3815, a_3 = -2.0159, a_4 = 10.3880, \theta = -0.0211$	22.33	2.88
	Average		16.25	2.10
Polynomial	Paper-cardboard	$-0.2483x_{income}^2 - 1.1596x_{income} + 6.5099 + 10.4326x_{population}^2 - 1.9047x_{income}x_{population} + 7.4633x_{population} - 11.2928x_{income}^2x_{population}^2 + 8.5430x_{income}x_{population}^2 - 0.7739x_{population}^2$	16.03	2.07
	Metal	$-1.8886x_{income}^2 - 0.7416x_{income} + 1.9012 - 1.4758x_{population}^2 - 0.5177x_{income}x_{population} - 0.1538x_{population} + 1.3301x_{income}^2x_{population}^2 - 0.9762x_{income}x_{population}^2 + 0.2854x_{population}^2$	9.02	1.16
	Glass	$-0.2369x_{income}^2 + 0.2126x_{income} + 2.2800 + 6.2291x_{population}^2 - 2.4592x_{income}x_{population} + 0.2447x_{population} - 9.3011x_{income}^2x_{population}^2 + 6.4913x_{income}x_{population}^2 - 0.9625x_{population}^2$	10.84	1.40
	Plastic	$-0.3208x_{income}^2 + 1.8168x_{income} + 3.7079 + 10.1965x_{population}^2 - 4.8257x_{income}x_{population} + 4.0275x_{population} - 14.7275x_{income}^2x_{population}^2 + 9.7465x_{income}x_{population}^2 + 3.8824x_{population}^2$	13.28	1.71
	Food waste	$3.6811x_{income}^2 + 1.7425x_{income} + 67.1681 - 54.7945x_{population}^2 + 41.3974x_{income}x_{population} - 8.0444x_{population} + 60.9312x_{income}^2x_{population}^2 - 53.3423x_{income}x_{population}^2 - 0.7690x_{population}^2$	23.90	3.09
	Others	$-7.4726x_{income}^2 + 2.1879x_{income} + 16.8696 + 27.0156x_{population}^2 - 34.2527x_{income}x_{population} + 1.7479x_{population} - 28.6273x_{income}^2x_{population}^2 + 27.5040x_{income}x_{population}^2 - 5.7457x_{population}^2$	22.00	2.84
	Average		15.85	2.04

5 Author contribution statements

In the scope of this study, Aysun ÖZKAN in the formation of the idea, the design and the literature review; Kemal ÖZKAN in the assessment of obtained results; Şahin IŞIK in formation of the models; Müfide BANAR in the editing the article in terms of content were contributed.

6 Ethics committee approval and conflict of interest statement

There is no need to obtain permission from the ethics committee for the article prepared. There is no conflict of interest with any person/institution in the article prepared.

7 References

[1] Chung SS. "Projecting municipal solid waste: The case of Hong Kong SAR". *Resources Conservation and Recycling*, 54(11), 759-768, 2010.

[2] Box GE, Jenkins GM. *Time Series Analysis: Forecasting and Control.* 5th ed. Holden-Day, USA, Wiley, 2015.

[3] Bridgwater A. "Refuse composition projections and recycling technology". *Resources Conservation and Recycling*, 12(3-4), 159-174, 1986.

[4] Mwenda A, Kuznetsov D, Mirau S. "Time series forecasting of solid waste generation in Arusha City-Tanzania". *Mathematical Theory and Modelling*, 4(8), 29-39, 2014.

[5] Sodanil M, Chatthong P. "Artificial neural network-based time series analysis forecasting for the amount of solid waste in Bangkok". *Digital Information Management (ICDIM) 2014 Ninth International Conference on. IEEE*, Phitsanulok, Thailand, 29 September-1 October 2014.

[6] Navarro-Esbrı J, Diamadopoulos E, Ginestar D. "Time series analysis and forecasting techniques for municipal solid waste management". *Resources Conservation and Recycling*, 35(3), 201-214, 2002.

[7] Bach H, Mild A, Natter M, Weber A. "Combining socio-demographic and logistic factors to explain the generation and collection of waste paper". *Resources Conservation and Recycling*, 41(1), 65-73, 2004.

[8] Ying L, Weiran L, Jingyi L. "Forecast the output of municipal solid waste in Beijing satellite towns by combination models". *Electric Technology and Civil Engineering (ICETCE) 2011 International Conference on IEEE*, Lushan, China, 22-24 April 2011.

[9] Xiang Z, Daoliang L. "Forecasting municipal solid waste generation based on grey fuzzy dynamic modeling". *IASME/WSEAS International Conference on Energy, Environment, Ecosystems and Sustainable Development*, Agios Nikolaos, Greece, 24-26 July 2007.

[10] Dai C, Li Y, Huang G. "A two-stage support-vector-regression optimization model for municipal solid waste management-a case study of Beijing, China". *Journal of Environmental Management*, 92(12), 3023-3037, 2011.

[11] Intharathirat R, Salam PA, Kumar S, Untong A. "Forecasting of municipal solid waste quantity in a developing country using multivariate grey models". *Waste Management*, 39, 3-14, 2015.

[12] Denafas G, Ruzgas T, Martuzevičius D, Shmarin S, Hoffmann M, Mykhaylenko V, Ogorodnik S, Romanov M, Neguliaeva E, Chusov A. "Seasonal variation of municipal solid waste generation and composition in four East European cities". *Resources Conservation and Recycling*, 89, 22-30, 2014.

[13] Zade MJG, Noori R. "Prediction of municipal solid waste generation by use of artificial neural network: A case study of Mashhad". *International Journal of Environmental Research*, 2(1), 13-22, 2007.

[14] Kumar JS, Subbaiah KV, Rao P. "Prediction of municipal solid waste with RBF net work-a case study of Eluru, AP, India". *International Journal of Innovation, Management and Technology*, 2(3), 238-243, 2011.

[15] Patel V, Meka S. "Forecasting of municipal solid waste generation for medium scale towns located in the state of Gujarat, India". *International Journal of Innovative Research in Science, Engineering and Technology*, 2(9), 4707-4716, 2013.

[16] Abbasi M, Abduli M, Omidvar B, Baghvand A. "Forecasting municipal solid waste generation by hybrid support vector machine and partial least square model". *International Journal of Environmental Research*, 7(1), 27-38, 2012.

[17] Özkan K, Işık Ş, Özkan A, Banar M. "Prediction for the solid waste composition by use of different curve fitting models: A case study". *Innovations in Intelligent SysTems and Applications (INISTA) 2015 International Symposium on IEEE*, Madrid, Spain, 2-4 September 2015.

[18] Onel O, Niziolek AM, Hasan MF, Floudas CA. "Municipal solid waste to liquid transportation fuels-Part I: Mathematical modeling of a municipal solid waste gasifier". *Computers and Chemical Engineering*, 71(4), 636-647, 2014.

[19] Yang N, Damgaard A, Kjeldsen P, Shao LM, He PJ. "Quantification of regional leachate variance from municipal solid waste landfills in China". *Waste Management*, 46, 362-372, 2015.

[20] Chapra SC, Canale RP. *Numerical Methods for Engineering*. United Kingdom, 7th ed. Tata McGraw Hill Education, 2015.

[21] Chai T, Draxler R. "Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature". *Geoscientific Model Development*, 7(3), 1247-1250, 2014.

[22] Banar M, Özkan A. "Characterization of the municipal solid waste in Eskisehir City, Turkey". *Environmental Engineering Science*, 25(8), 1213-1220, 2008.

[23] Özkan A. Utilization of Different Decision Making Techniques for Development of Municipal Solid Waste Management Systems. pHD Thesis, Anadolu University, Eskisehir, Turkey, 2008.

[24] Jalili Ghazi Zadeh M, Noori R "Prediction of municipal solid waste generation by use of artificial neural network: A case study of Mashhad". *International Journal of Environmental Research*, 2(1), 13-22, 2008.

[25] Kannangara M, Dua R, Ahmadi L, Bensebaa F "Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches". *Waste Management,* 74, 3-15, 2018.

[26] Araiza-Aguilar JA, Rojas-Valencia MN, Aguilar-Vera RA "Forecast generation model of municipal solid waste using multiple linear regression". *Global Journal of Environmental Science and Management*, 6(1), 1-14, 2020.