Waste Management 48 (2016) 14–23

Contents lists available at ScienceDirect

Waste Management

journal homepage: www.elsevier.com/locate/wasman

Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars province, Iran

Sama Azadi, Ayoub Karimi-Jashni *

Department of Civil and Environmental Engineering, School of Engineering, Shiraz University, Shiraz, Fars 71348-51156, Iran

A R T I C L E I N F O

Article history:
Received 10 April 2015
Revised 18 July 2015
Accepted 24 September 2015
Available online 9 October 2015

Keywords:
Artificial neural network
Seasonal municipal solid waste generation
Fars Province
Multiple linear regression

A B S T R A C T

Predicting the mass of solid waste generation plays an important role in integrated solid waste management plans. In this study, the performance of two predictive models, Artificial Neural Network (ANN) and Multiple Linear Regression (MLR) was verified to predict mean Seasonal Municipal Solid Waste Generation (SMSWG) rate. The accuracy of the proposed models is illustrated through a case study of 20 cities located in Fars Province, Iran. Four performance measures, MAE, MAPE, RMSE and R were used to evaluate the performance of these models. The MLR, as a conventional model, showed poor prediction performance. On the other hand, the results indicated that the ANN model, as a non-linear model, has a higher predictive accuracy when it comes to prediction of the mean SMSWG rate. As a result, in order to develop a more cost-effective strategy for waste management in the future, the ANN model could be used to predict the mean SMSWG rate.

1. Introduction

The Municipal Solid Waste (MSW) – more commonly known as trash or garbage – is one of the most challenging problems in modern societies, especially in many developing countries. As solid waste management is a relatively new subject in developing countries, in order to develop and implement effective solid waste management programs it is necessary to have an accurate and reliable prediction of the mass of Municipal Solid Waste Generation (MSWG) (Batinic et al., 2011; Kumar et al., 2011; Cherian and Jacob, 2012; Intharathirat et al., 2015). The quantity of generated and collected solid wastes is of great importance in the solid waste management program (Tchobanoglous et al., 1993; Fu et al., 2015).

Solid Waste Generation (SWG) rate is affected by a variety of factors including the geographical location, season, use of kitchen food waste grinders, collection frequency, characteristics of service area, onsite processing, people's food habits, economic conditions, recovery and reuse boundaries, existence of solid waste management laws, local culture and beliefs, population growth, weather conditions and household size (Tchobanoglous et al., 1993; Abdoli et al., 2011; Chen, 2010; Keser et al., 2012; Intharathirat et al., 2015; Mateu et al., 2013; Lebersorger and Beigl, 2011; Boer et al., 2010; Boldrin and Christensen, 2010). Gomez et al. (2009) studied MSWG in three neighborhoods of Chihuahua, Mexico with different income levels in three seasons of year. Their results showed that higher income generally leads to the greater amounts of SWG and less waste generated during winter.

An integrated solid waste management system requires the precise prediction of SWG rate. One common approach in predicting SWG is the consideration of the trend of waste generation rate variation along with time in a target area. Unfortunately, there is a lack of comprehensive historical records on the SWG rate in many developing countries due to insufficient funds, inadequate management and limited measurement instruments in many cases (Rimaityte et al., 2011; Dyson and Chang, 2005). Thus, the prediction of the SWG rate by common methods, which consider the trend of the waste generation rate variation along with its time, is quite challenging. In order to tackle this problem effectively, it is suggested that a new approach be developed which does not require the trend of waste generation rate variation along with its time to predict the SWG rate.

Various studies have made use of many forecasting techniques such as multiple linear and nonlinear regression techniques and expert systems to simulate SWG (Jalili Ghazi Zade and Noori, 2008; Kollikkathara et al., 2010; Dyson and Chang, 2005; Lohani

* Corresponding author.
E-mail address: akarimi@shirazu.ac.ir (A. Karimi-Jashni).

http://dx.doi.org/10.1016/j.wasman.2015.09.034
0956-053X/

and Hartono, 1985). Lohani and Hartono (1985) formulated several equations for determining SWG rates in different domestic and industrial sources in the city of Bandung in Indonesia using MLR model. The authors concluded that the presented formulas were useful for the prediction of SWG rates in Bandung. Dyson and Chang (2005) used a dynamic system modeling technique to predict the mass of MSWG in San Antonio, Texas. They concluded that dynamic modeling can be used when there is a lack of complete historical records on SWG, and traditional statistical least-squares regression methods are not applicable. Jahandideh et al. (2009) used ANN and MLR models to predict the medical waste generation rate of 50 hospitals in Fars Province, Iran. The results showed that ANN could predict SWG rate reasonably well, while the prediction achieved by the MLR model was not adequately accurate. Jalili Ghazi Zade and Noori (2008) investigated different structures of feed forward ANN model to predict the SWG volume. They concluded that a network with 16 neurons in the hidden layer was suitable for the prediction of the weekly volume of waste generated in Mashhad, Iran. Batinic et al. (2011) used ANN model to predict the amount and composition of SWG in Serbia. They indicated that ANN model was capable and suitably effective in this assignment. Abdoli et al. (2011) applied ANN and MLR models to a long-term (over 22 years) prediction of SWG mass by analyzing monthly series of datasets over the first decade of the 21st century (2000–2010) in Mashhad, Iran. The results showed that ANN can follow trends and fluctuations more accurately than MLR model.

It is worth mentioning that MLR techniques, which are widely applied in forecasting MSWG, are mainly used to model a linear relationship between a dependent variable and one or more independent variables (Xu et al., 2013; Thanh et al., 2010). Thus, non-linear and dynamic modeling techniques such as fuzzy systems, evolutionary programming, ANN and the combination of these methods are essential for the development of accurate predictive models (Intharathirat et al., 2015; Patel and Meka, 2013).

Interest in using ANNs for forecasting has increased in research activities during the past decade (Patel and Meka, 2013; Roy et al., 2013; Shahabi et al., 2012; Jalili Ghazi Zade and Noori, 2008; Batinic et al., 2011). The ANNs are a class of non-linear mathematical models which were developed on the basis of human neural system. The ANN models are characterized by architectures and structures of interconnected computational elements, i.e. neurons. These architectures enable ANN to learn from examples and capture functional relationships among the data and can often correctly infer the unseen part of sample data even if the data contain some noise (Zhang et al., 1998).

In this study, two predictive models ANN and MLR were applied to verify the performance of the models for the prediction of mean SMSWG rate in 20 urban regions of Fars Province, Iran. Five influential variables on the mean SMSWG rate, population, solid waste collection frequency, maximum seasonal temperature and altitude were chosen as the input data of our study based on correlation test.

2. Material and methods

2.1. Model framework

Fig. 1 shows the main components of the proposed methodology to predict the mean SMSWG rate. As shown in this figure, the methodology includes three main parts: (1) specifying explanatory variables and data collection, (2) developing ANN and MLR models, training and testing them, (3) statistically analyzing the results and determining the most accurate predictive model. In the first step, effective parameters for the simulation of the mean SMSWG rate were investigated, and some measurable parameters and those whose data were available were selected. In the second step, ANN and MLR models were developed. In this step, the most suitable structure of ANN model for the mean SMSWG rate prediction was specified. In addition, before developing MLR model, principal component analysis was conducted on the explanatory variables to remove their multicollinearity. In the third step, the performances of the applied models were compared, and based on statistical tests, the most accurate model for the prediction of the mean SMSWG rate was determined. More details on the components of the proposed methodology will be discussed in the following sections.

2.2. The study area and predictor variables

Fars Province is located in the southwest of Iran from 27°03′ and 31°42′N and 50°30′ and 55°36′E (see Fig. 2). Fars Province has an area of about 122,608 km^2, covering approximately 8.1% of Iran (Nafarzadegan et al., 2010). According to the latest figures provided by the Census Bureau of Iran (2011), it has a population of about 4.58 million, constituting about 6.15% of the total population of Iran. About 61.2% and 38.1% of the population of Fars live in urban and rural areas, respectively, and about 0.7% are nomads. Due to the vastness of the province and the presence of the Zagros Mountains to the north and northwest, the plain of Khuzestan to the west, and the Persian Gulf to the south, the climate and weather conditions are highly variable. The difference in average annual rainfall between the northern and southern parts of the province exceeds 300 mm.

The changes of the mean SMSWG rate in spring, summer, autumn and winter of 2009–2010 for 20 cities in Fars Province are shown in Fig. 3. As can be seen, the mean SMSWG rate in all cities approximately follows the same trend within the year and it can be said that the per capita generation of MSW increases in summer which is due to the dietary habits of the people, the rise in tourist trips, and the availability of more fruits and vegetables. However, the appropriate measures can be taken by predicting the mean SMSWG rate in order to determine the frequency of waste collection and estimate the number of required trucks to carry the waste in different seasons and finally reduce the costs.

In this study, based on the correlation test and previous related literature, four influential variables on SMSWG rate whose data were available were chosen as the explanatory variables of our study. The selected variables were required to be able to be forecasted with a relatively high accuracy in a specified forecasting horizon.

The first selected variable was the population of each city because the consumption pattern changes as the population grows and therefore the per capita generation of MSW increases (Abdoli et al., 2011; Dyson and Chang, 2005). The second variable considered as an explanatory variable was the frequency of waste collection, because if the frequency of waste collection in a city increases, the collected waste will increase (Tchobanoglous et al., 1993; Keser et al., 2012). The third selected parameter was the maximum seasonal temperature. This parameter which was previously utilized as an effective parameter on SWG rate by Abdoli et al. (2011) was used to show the change of seasons. On the other hand, one of the parameters that affect the solid waste generation rate in different cities is the climate of each city. So, the fourth selected variable was the altitude of each city because this parameter can be a sign of changing weather conditions as well as the geographic and climatic conditions of each city.

Hence, the four explanatory variables population (Pop), solid waste collection frequency (Freq), maximum seasonal temperature (T_{max}) and altitude (Alt) were considered for the prediction of the mean SMSWG rate. Thus, the 2009–2010 data for 20 urban areas of Fars Province which had air station were used in this study. Table 1 indicates the statistical characteristics of the data. Also,