ARTIFICIAL NEURAL NETWORK APPLIED IN FORECASTING THE COMPOSITION OF MUNICIPAL SOLID WASTE IN IASI, ROMANIA

Cristina GHINEA1, 2*, Petronela COZMA2, Maria GAVRILESCU2*

1Faculty of Food Engineering, “Stefan cel Mare” University of Suceava, 13 Universitatii Street, 720229 Suceava, Romania
2Department of Environmental Engineering and Management, “Cristofor Simionescu” Faculty of Chemical Engineering and Environmental Protection, “Gheorghe Asachi” Technical University of Iasi, 73 Prof. D. Mangeron Blvd., 700050 Iasi, Romania

Highlights

- Significant factors for planning, development of sustainable MSW management system.
- Neural network time series used for prediction of municipal solid waste composition.
- Two ANN models (NIO and NARX) were applied as NNTS tools.
- The coefficient of determination (R2) and root mean square error (RMSE) were used.
- Information for decision-makers to improve the management system of MSW provided.

Abstract. Neural network time series (NNTS) tool was used to predict municipal solid waste composition in Iasi, Romania. The nonlinear input output (NIO) time series model and nonlinear autoregressive model with external (exogenous) input (NARX) included in this tool were selected. The coefficient of determination (R2) and root mean square error (RMSE) were chosen for evaluation. By applying NIO, the optimum model is 4-11-6 artificial neural network (ANN, R2 = 0.929) in the case of testing as for the validation, with 0.849 and 0.885, respectively. Applying NARX, the suitable model became 4-13-6 ANN model, with R2 = 0.999 for training, 0.879 for testing, and 0.931, respectively 0.944 for validation and all. The resulted RMSE is zero for training and 0.0109 for validation in the case of this model which had 4 inputs, 13 neurons and 6 outputs. The four input variables were: number of residents, population aged 15–59 years, urban life expectancy, total municipal solid waste (ton/year). The suitable ANN model revealed the lowest root mean square error and the highest coefficient of determination. Results indicate that NNTS tool is a complex instrument, NARX is more accurate than NIO model, and can be used and applied easily.

Keywords: artificial neural network, environmental processes modeling, population, solid waste, urban life expectancy, waste composition, waste management technologies.

Introduction

There are many significant factors which influence planning and development of sustainable solid waste management systems and one of them is solid waste generation (Ghinea et al., 2016; Henriques & Coelho, 2019; Li et al., 2016; Ortiz-Rodriguez et al., 2018; Simion et al., 2012). Prediction of solid waste (SW) quantities that will be generated represents a challenge for all the decision makers involved in the waste management (Ghinea & Gavrilescu, 2010, 2019; Younes et al., 2015). Different tools were developed and proposed for solid waste prediction using quantitative techniques based on mathematical models: time series models (Ghinea et al., 2016; Owusu-Sekyere et al., 2013; Rimaityte et al., 2012), artificial neural network (ANN) (Adamović et al., 2016; Batinic et al., 2011; Radaei et al., 2017; Younes et al., 2015), multiple regression analysis (Kumar & Samadder, 2017; Shan, 2010) and others. Artificial neural network (ANN) represents one of the most applied tools (Abbasi & El Hanandeh, 2016; Abbasi et al., 2014; Goel et al., 2017; Kannangara

*Corresponding author. E-mails: mgav@tuiai.ro; cristina.ghinea@fia.usv.ro

Copyright © 2021 The Author(s). Published by Vilnius Gediminas Technical University

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
An ANN model was proposed by Antanasijević et al. (2013) for prediction of waste generated in 44 countries considering factors like: countries size, inhabitant's number, economic and social aspects and others for the period from 2000 to 2012. In order to determine the solid waste amounts that will be generated, Younes et al. (2015) considered the input variables such as (Table 1): number of inhabitants and also employed/unemployed persons, gross domestic product (GDP) and energy demand. They used nonlinear autoregressive with external (exogenous) input model (NARX) for prediction (Younes et al., 2015). In another study the authors used data from 2004 to 2009 and applied ANN and response surface model for solid waste generation forecasting (Shamshiry et al., 2014). The input variables considered by Shamshiry et al. (2014) were: fuel consumption, 4-ton truck, 10-ton truck, number of trips made by the trucks to the landfill, number of times the personnel entered into the landfill, number of tourists, and salary per worker per day.

The ANN was also applied for waste prediction in Serbia by Batic et al. (2011). The inputs variables were: municipality income, employment/unemployment, age, educational and social aspects, while the output variables were six different waste fractions (Table 1). The suitable ANN model in this case was 4-10-6: 4 inputs, 10 hidden layers and 6 outputs (Batic et al., 2011). ANN modelling was also chosen by Kumar et al. (2011) for prediction of waste in one city from India for the period 2010–2026 considering inputs such as: population number, percentage of urban population, municipal solid waste generated, GDP. Prediction of solid waste from 26 European countries was investigated by Antanasijević et al. (2013) developing ANN models. Waste generation weight was predicted by Zade and Noori (2008) during different seasons, over three years using also ANN. Their results showed that after the investigation of 12 model structures the suitable are structures with 10 and 16 neurons in the hidden layer (13-10-1 and 13-16-1). They selected from these models the structure 13-16-1 which seems to be the most appropriate for waste prediction based on mean absolute error, mean absolute relative error, root mean square error, correlation coefficient and threshold statistics.

ANN models are using previous data for prediction, which makes them suitable for forecasting. They have the ability to plot input and output (Ali & Ahmad, 2019). The nonlinear relationship description and time series simulation can be performed with NARX model which is a ANN prediction model used in many different studies (for prediction of waste generation, solar radiation, gaseous emissions from wastewater plants, electricity prices etc.) (Boussaada et al., 2018; Marcjasz et al., 2019; Vu et al., 2019; Zounemat-Kermani et al., 2019). According to Boussaada et al. (2018), Yu et al. (2019) NARX neural network is a good predictor of time series, widely used and well-fitting. In addition, the novelty of our work consists in the way of applying the ANN methodology for the forecast of solid waste composition and especially by comparing the results obtained using two ANN models (NARX and NIO). Our study provides information on the application of ANN for selection of social and demographic indicators which influence the estimation of solid waste and also for establishing the suitable ANN structure in order to predict the amounts of various waste fractions that will be generated.

Therefore, there is a perpetual need for waste management planning. Forecasting is the first step which has to be performed before designing and implementation of municipal solid waste management systems. It is important the selection of the most suitable model for forecasting in order to obtain the most accurate results. The $R^2$ values normally ranges from 0 to 1. Good predictions are

| Method                  | Input variables                                                                 | Output variables                  | Studies                          |
|-------------------------|---------------------------------------------------------------------------------|-----------------------------------|----------------------------------|
| ANN model               | number of trucks which carry waste                                             | waste generation                  | Zade and Noori (2008)            |
| ANN model               | average household income                                                       | organic waste                     | Batic et al. (2011)              |
|                         | percentage of employees on municipality level                                   | paper and cardboard               |                                  |
|                         | average age of population                                                       | plastic                           |                                  |
|                         | level of education                                                             | glass                              |                                  |
|                         | household size                                                                 | metal                              |                                  |
| Multiple linear regression technique | total family income  | non-biodegradable waste          | Kumar et al. (2011)              |
| AN model General regression neural network | gross domestic product per capita | municipal waste generation          | Antanasijević et al. (2013)    |
| AN model and response surface model | personnel number  | solid waste                        | Shamshiry et al. (2014)          |
| AN model (NARX)         | salary per worker per day                                                       | paper waste                        | Younes et al. (2015)             |
|                         | fuel consumption                                                                | plastic                            |                                  |
|                         | lorry number                                                                    | glass                              |                                  |
|                         | number of tourists                                                              | metal                              |                                  |
| AN model (NIO and NARX) | number of residents                                                            | biodegradable waste               | this study                       |
|                         | population aged 15 to 59 years                                                  | -                                 |                                  |
|                         | urban life expectancy                                                           | -                                 |                                  |
|                         | total municipal solid waste                                                     | -                                 |                                  |

Table 1. Input and output variables used for prediction of waste generation
performed when $R^2$ is very close to 1 and worse predictions when $R^2$ is 0 or very close to this value. The $R^2$ values obtained in different studies were: $R^2 = 0.97$ (Younes et al., 2015), $R^2 = 0.82$ (Singh & Satija, 2016), $R^2 = 0.81$ (Zade & Noori, 2008) and $R^2 = 0.64$ (Shahabi et al., 2012). The suitable ANN model is considered the model that has the lowest root mean square error and the highest regression coefficient ($R^2$).

In this context, the aim of this paper is to predict the amount of municipal solid waste fractions by using ANN models, to investigate the influence of social variables on waste generation and to compare the results obtained by applying two ANN models (NIO and NARX).

1. Materials and methods

1.1. Case study

Iasi is a city belonging to the Iasi County, which is a part of the Region 1 North East (NE) established by Law no. 315/2004 on regional development in Romania. The Romanian National Strategy for Waste Management and the National Plan for Waste Management are the basis for waste management in Romania (Romanian National Plan Waste [RNPWM], 2018; Romanian National Strategy Waste, 2013). According to the National Waste Management Plan and the Regional Management Plan for Region NE the responsibility for the municipal waste belongs to the local public authorities and/or to an authorized economic operator.

The prediction is performed in order to provide and apply a modelling tool to help the decision and policy makers and stakeholders concerned about waste management. For this purpose, data regarding the number of inhabitants, population aged 15 to 59 years, urban life expectancy and amounts of municipal solid waste generated for the following years: 2000–2018 were collected from Iasi County Statistics, Iasi County Council and previous studies and these data were used for training and validation. The forecasted data for 2019–2025 were also considered for testing. The data used are at municipality level. The selection of these variables has been performed based on other studies which demonstrated that these variables were linked with waste generation and also they had a significant influence on waste prediction than others (Ghinea et al., 2016; Rimaityte et al., 2012).

The number of inhabitants of Iasi City increased from 410523 in 2000 to 435887 in 2015 (Figure 1a) according to the National Institute of Statistics data (Ghinea et al., 2016; National Institute of Statistics [NIS], 2015), followed by a slight decrease in 2016 and then grow again in 2018 at 430687 inhabitants. The population aged 15 to 59 years’ indicator for the period 2000–2018 varied between 60 and 73% (Figure 1b), with the highest value registered in 2006 (72.78%). The urban life expectancy increased in the last years, the values for this indicator increased from 68 to 76.48 years (Figure 1c) with the highest value registered in 2018. The municipal solid waste quantity also increased (Figure 1d) mainly due to population higher consumption and a higher percentage of the population served by public health services in a centralized system.

The fractions identified in Iasi, Romania and selected for this modelling were: paper and cardboard, plastic, glass, metals, biodegradable waste and others (Figure 2). The fraction with the major percentage in the composition

![Figure 1. Input variables: a) $x_1$ – number of inhabitants; b) $x_2$ – population aged 15 to 59 years (%); c) $x_3$ – urban life expectancy (years); d) $x_4$ – municipal solid waste amount (t/year)](image)
of solid waste generated by population is biodegradable waste (approx. 50%) (Figure 2) followed by paper and cardboard, plastic, glass, metals and other waste (Ghinea et al., 2016; Ghinea, 2012).

The municipal solid waste generated is collected by a public company and efforts are made for separate waste collection. Several Romanian companies, in the last years, conducted waste recycling campaigns to involve the inhabitants in the selective collection process. In Romania, waste separate collection is performed usually for materials with high market values, mainly in pilot projects (Ghinea & Gavrilescu, 2019). The municipal solid waste management system in Iasi is presented in our previous works (Ghinea et al., 2016, 2012; Ghinea & Gavrilescu, 2016).

Ghinea et al. (2016) applied the Waste Prognostic Tool, regression analysis and time series analysis for prediction of municipal solid waste generation in Iasi city. The same inputs data have been used in this paper to predict SW fractions generation by applying ANN modeling. The neural network time series (NNTS) tool was chosen for the modeling and evaluation.

1.2. ANN methodology

ANN is a model that belongs to artificial intelligence. It consists of artificial neurons connected with coefficients, thus forming the neural structure. Compared to the human brain, this is far behind, but can be used to process a large amount of data and to make predictions. In the ANN model, the artificial neuron simulates the function of the biological neuron. Inputs represent the signals that come and are activated, combined, and usually with the help of the transfer function (mainly the sigmoid function) the outputs are produced (Agatonovic-Kustrin & Beresford, 2000). These outputs are generated when the inputs exceed a certain threshold (Han et al., 2018). Therefore, the ANN is a mathematical tool inspired by biological neural networks which includes input, output and hidden layers (Adamović et al., 2016). The inputs and outputs represent variables which will be introduced into the model, while the hidden neurons are determined by trial and error (Younes et al., 2015). Inputs data will be received by input layer neurons, while the response to the input data are provided by the output neurons (Agatonovic-Kustrin & Beresford, 2000; Han et al., 2018). Neurons from the hidden layers are connected only with others neurons, and a single hidden layer is sufficient according to Agatonovic-Kustrin and Beresford (2000). In the learning process, ANN use a weight modification method in order to adjust weights just to improve outputs. During prediction, fed backwards through the network is performed for the adjustment of weight and error minimization. This prevents the same error from occurring again. In the first phase (training), inputs are transformed by the connection weights. Number of neurons is particularly important in this phase, because too few neurons will hinder the learning process, and too many will cause an overload. The

![Figure 2. Output variables: a) paper waste; b) plastic waste; c) glass waste; d) metal waste; e) biodegradable waste; f) other waste (all in %)]
weighted are used to predict data only after the desired outputs are produced. In the prediction phase network works only by forward propagation of data and the output of this propagation represents the prediction model for data validation (Agatonovic-Kustrin & Beresford, 2000). The ANN modeling can be performed with NNTS tool which is an instrument that can be applied to solve three nonlinear time series problems. The NNTS tool includes: nonlinear autoregressive model with external (exogenous) input (NARX), nonlinear autoregressive model (NAR) and nonlinear input output model (NIO) (MathWorks, 2016).

The first technique used was nonlinear input output (NIO), in which the series \( y(t) \) is predicted based on the predicted series (Eq. (1)) (MathWorks, 2016):

\[
y(t) = f(x(t-1), x(t-2), \ldots, x(t-d)) .
\]

The second one technique applied was NARX, which is a dynamic feedforward network with backpropagation (lag) and is a multiple-layer network (Younes et al., 2015). The series \( y(t) \) is predicted based on the following equation (Eq. (2)) (MathWorks, 2016):

\[
y_t = f(x(t-1), \ldots, x(t-d), \ldots, y(t-1), \ldots, y(t-d)) ,
\]

where: \( y(t) \) series given \( d \) past values of \( y(t) \) and another series \( x(t) \), \( d \) is tapped delay lines that store the previous values of \( y(t) \) and \( x(t) \) sequences.

Two statistical indices were used to evaluate the performance of ANN models:
- the root mean squared error (RMSE) (Eq. (3)):

\[
RMSE = \sqrt{ \frac{\sum_{i=1}^{n} (x_i - x_0)^2}{n} } .
\]

RMSE is a performance function and measures the network’s performance according to the mean of squared errors and also can be defined as the average squared differences between outputs and targets.
- the coefficient of determination \( R^2 \) (Eq. (4)):

\[
R^2 = \frac{\sum (x_0 - x_i)^2}{\sum (x_0 - \overline{x})^2 + \sum (x_i - \overline{x})^2} ,
\]

where \( x_i \) is the actual output and \( x_0 \) is the predicted output, \( n \) is the number of the outputs, \( \overline{x} \) is the mean of actual output and \( \overline{x_0} \) is the mean of the predicted output.

The \( R^2 \) values represent the correlation between output and targets and there is a close connection between these two when \( R^2 = 1 \).

\( R^2 \) and RMSE indices are the most common used in waste prediction (Abbas & Hanandeh, 2016; Adamović et al., 2016; Kannanagara et al., 2018; Shahabi et al., 2012; Younes et al., 2015; Zade & Noori, 2008 and others). The default metric for many models is still RMSE, due to the fact that the loss function defined in terms of RMSE is easy to differentiate and the mathematical operations are performed much easier. \( R^2 \) shows information about how well the variation in the output was explained by the targets and outputs.

This tool also includes different types of algorithms such as: (1) Levenberg-Marquardt (LM) backpropagation; (2) Bayesian Regularization (BR) algorithm and (3) Scaled Conjugated Gradient (SCG) (Noori et al., 2010b, 2011).

The most used is Levenberg-Marquardt (LM) algorithm (Ranganathan, 2004) which is considered the fastest training function as well (MathWorks, 2016). The LM method description can be found in literature (Ranganathan, 2004). Different authors used LM for prediction of municipal solid waste generation in Malaysia and India respectively (Singh & Satija, 2016; Younes et al., 2015). The LM algorithm was also selected for application in this study.

The NNTS tool from Matlab version 9.1 (MathWorks, 2016) was chosen in this study for prediction of SW fractions generation. Figure 3 illustrates the network modeling steps followed by us.

NIO and NARX were applied in order to establish the suitable model and to compare them with each other. For this study the input variables are: \( x_1 \) – number of residents, \( x_2 \) – population aged 15 to 59 years, \( x_3 \) – urban life expectancy and \( x_4 \) – total municipal solid waste (ton/year). The inputs correspond to those presented in the previous work (Ghinea et al., 2016) for the period 2000–2015, while for 2016–2018 were taken from Iasi County Statistics and Iasi County Council and the forecasted data (2019–2025) from (RNPWM, 2018) and Iasi County Statistics. Inputs data is a 26×4 matrix representing dynamic data 26 time steps of 4 elements \( (x_1, x_2, x_3, x_4) \), while targets is a 26×6 matrix representing dynamic data 26 time steps of 6 elements (paper waste, plastic waste, glass waste, metal waste, biodegradable waste and other waste). Figures 1 and 2 illustrates the values of input and output variables considered in this research. All data were normalised by using Eq. (5) (Zade & Noori, 2008):

\[
X_{\text{norm}} = 0.8 \left( \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \right) + 0.1 .
\]

![Figure 3. The network modelling phases](image-url)

Figure 3. The network modelling phases
The model architecture used is presented in Figure 4. The training data were set at 55% for training, 20% for validating and 25% for testing, respectively. A similar procedure was also chosen and applied by Singh and Satija (2016). The network generalization is represented by the validation time steps, while the network performance is shown through time steps testing. The training is considered complete when the generalization stops improving.

Data conditioning are: for training were used the data from 2000–2013, for validation the data from 2014–2018 and for testing phase the data obtained for 2019–2025 (predicted data).

In this study it was also checked if there are any outliers. According to NIST/SEMATECH (2015) the outliers are telling us something and in this case there is recommended to graphically check outliers by using a box plot or quantitatively check for outliers by Grubbs test for a single outlier, while for more than one outlier Tietjen-Moore test must be performed. It can be established how far the outlier is from the others by calculus. $Z$ score is used for the detection of outliers (Guthrie, 2008) and is determined with Eq. (6):

$$Z = \frac{Y_i - \overline{Y}}{SD},$$ (6)

where: $Y_i$ – sample, $\overline{Y}$ – sample mean, $SD$ – sample standard deviation.

$Z$ score determination is important, since it represents a way for results comparison with a normal (Gaussian) distribution and also enables comparison of two scores from different samples. In this paper, $Z$ score was determined by using Excel file.

2. Results and discussion

In order to establish the most suitable network structure for prediction of SW fractions generation, several neural networks structure with input, hidden and output layers were investigated. Different neurons in the hidden layer were considered. In the first part of the evaluation of NIO model was considered, while in the second part NARX model was assessed.

2.1. Prediction of SW fractions by using NIO model

Table 1 presents the ANN models structure, root mean squared error (RMSE) and coefficient of determination ($R^2$) values for training ($tr$), testing ($ts$), validation ($v$) and all ($a$). For the ANN model 4-11-6 the $R^2$ = 0.883 in the case of $tr$ and $R^2$ = 0.929 for $ts$, while $R^2$ it is 0.849 and 0.885 for the $v$ and $a$. It can be observed that the suitable ANN model with maximum correlation $R^2$ compared with the other ANN models is the 4-11-6 ANN model.

For RMSE the lower values are considered better, while a 0 value shows no error. From Table 2 it can be observed that

### Table 2. The results of training and testing stages of ANN using NIO model

| ANN model structure | RMSE      | Coefficient of determination ($R^2$) |
|---------------------|-----------|--------------------------------------|
|                     | Training ($tr$) | Testing ($ts$) | Validation ($v$) | Training ($tr$) | Testing ($ts$) | Validation ($v$) | All ($a$) |
| 4-4-6               | 0.0053    | 0.0298    | 0.0252    | 0.947 | 0.771 | 0.764 | 0.858 |
| 4-5-6               | 0.0053    | 0.0304    | 0.0371    | 0.955 | 0.770 | 0.619 | 0.833 |
| 4-6-6               | 0.0231    | 0.0165    | 0.0059    | 0.806 | 0.8744 | 0.925 | 0.841 |
| 4-7-6               | 0.0103    | 0.0246    | 0.0427    | 0.906 | 0.831 | 0.616 | 0.808 |
| 4-8-6               | 0.0037    | 0.0611    | 0.0547    | 0.967 | 0.546 | 0.510 | 0.756 |
| 4-9-6               | 0.0067    | 0.0965    | 0.0289    | 0.940 | 0.296 | 0.798 | 0.727 |
| 4-10-6              | 0.0397    | 0.0724    | 0.0741    | 0.964 | 0.641 | 0.545 | 0.741 |
| 4-11-6              | 0.0137    | 0.0089    | 0.0156    | 0.883 | 0.929 | 0.849 | 0.885 |
| 4-12-6              | 0.0095    | 0.0402    | 0.0368    | 0.916 | 0.651 | 0.593 | 0.795 |
| 4-13-6              | 0.0062    | 0.0372    | 0.0264    | 0.943 | 0.704 | 0.741 | 0.838 |
| 4-14-6              | 0.0130    | 0.0192    | 0.0611    | 0.895 | 0.845 | 0.493 | 0.802 |

Figure 4. Model architecture used for SW fractions prediction with NIO model
for all ANN model structures the RMSE values are close to 0. The lowest RMSE values for \(tr\), \(ts\) and \(v\) were recorded by considering different type of models (4-8-6, 4-11-6 and 4-6-6). The values presented in Table 2 are obtained by applying the Levenberg-Marquardt training algorithm.

Considering that the ANN model 4-11-6 is suitable, the results obtained by considering 11 hidden neurons in hidden layers and 2 as the number of time delays were discussed further.

It was observed that when the validation error increased for six iterations the training stops and this was appeared at iteration 8 (Figure 5a). At epoch two it was obtained the proper validation and it was observed that the next epochs did not reduce the errors. In Figure 5b the training data are represented by the blue bars, the validation data by green bars and the testing data by the red bars.

The histogram illustrated in Figure 5b is skewed right with one tail of the distribution considerably longer on the right side. It can be observed one outlier for training points. The Grubbs test was performed in order to check if this point is an outlier or not. For a significance level 0.05 (two-sided) with the critical value of \(Z = 2.680\) the point is furthest from the rest, but not a significant outlier (\(Z = 1.628\) for this point), which means that this point is not an outlier. Figure 6 shows the linear regression among the model outputs and the corresponding targets. The values of \(R^2\) present the relationship between observed and predicted values of solid waste during \(tr\), \(ts\) and \(v\) phases of the 4-11-6 ANN model. \(R^2\) values were calculated based on the difference between the predicted values and actual amount of SW fractions generated. According to Table 2 the best result is obtained using 4-11-6 ANN structure

![Figure 5. a) Best validation performance; b) error histogram](image)

![Figure 6. Regression values obtained for training, testing, validation and all for 4-11-6 model (target – observed values, output – predicted values)](image)
and is very good illustrated by Figure 6. The best fit linear regression line between outputs and targets is provided by the solid line (MathWorks, 2016). The equations are also illustrated in the Figure 6.

2.2. Prediction of SW fractions by using NARX model

Table 3 illustrates the ANN models structure, RMSE and $R^2$ values for $tr$, $ts$, $v$ and $a$ obtained by applying NARX. The Levenberg – Marquardt training algorithm was also considered. From Table 3 it can be notice that the ANN model 4-13-6 is the adequate model: since RMSE values are lower, and also the $R^2$ is equal with 0.999 for $tr$, 0.879 for $ts$, 0.931 for $v$ and 0.944 for $a$. NARX proposes the dependence between SW generation by fractions and inputs variables. In this study, we have chosen 13 hidden neurons in hidden layers and we selected the time delays 2.

From Figure 7a it can be observed that the training stopped when the validation error increased for 2 iterations. This was occurred at iteration 7. Also, the best validation appeared at epoch 5. Results showed that the performance gradient magnitude is $1.3455 \times 10^{-9}$, while the validation checks are 2 at 7 epochs (Figure 7b). From Figure 7c it can be observed one possible outlier for testing points.

The Grubbs test was performed and confirmed that for a significance level 0.05 (two-sided) with the critical value of $Z = 2.680$ the last point is not an outlier ($Z = 1.676$). In Figure 8 is presented the linear regression obtained between the outputs and targets data. It can be notice that the outputs closely follow the targets. The values of $R^2$ were calculated and presented in Figure 8 for $tr$, $ts$ and $v$ phases of the 4-13-6 ANN model.

Figure 9a illustrates the comparison between the actual and predicted SW fractions generation. There are five major errors illustrated by orange lines, one of them is for testing of one target point and the others for validation targets. The model is able to represent the targeted output.

In Figure 9b the autocorrelation of error 1 is illustrated, which describes how the prediction errors are related in time. In Figure 9b it is only one value that exceeds confidence limit and is at zero lag. This fact indicated a perfect prediction model and also that the prediction errors are completely uncorrelated with each other (MathWorks, 2016). From Figure 9b the model seems to be adequate since the correlations fall within the confidence limits.

Eleven neurons in the hidden layer are the adequate number when ANN model is 4-11-6 as can be seen in Figure 6. The following linear equation was obtained for 4-11-6 model while when 4-13-6 model was considered the tool provides the next equation which can be considered as suitable (Figure 8). In the other studies regarding

| ANN model structure | RMSE | Coefficient of determination ($R^2$) |
|---------------------|------|--------------------------------------|
|                     | Training ($tr$) | Testing ($ts$) | Validation ($v$) | Training ($tr$) | Testing ($ts$) | Validation ($v$) | All ($a$) |
| 4-4-6               | 0.0221 | 0.0220 | 0.0053 | 0.806 | 0.795 | 0.949 | 0.829 |
| 4-5-6               | 0.0163 | 0.0087 | 0.0523 | 0.889 | 0.940 | 0.517 | 0.840 |
| 4-6-6               | 0.0092 | 0.0212 | 0.0157 | 0.915 | 0.827 | 0.849 | 0.878 |
| 4-7-6               | 0.0065 | 0.0246 | 0.0229 | 0.943 | 0.787 | 0.856 | 0.880 |
| 4-8-6               | 0.0055 | 0.0159 | 0.0562 | 0.959 | 0.881 | 0.535 | 0.846 |
| 4-9-6               | 0.0033 | 0.0928 | 0.0140 | 0.981 | 0.630 | 0.949 | 0.831 |
| 4-10-6              | 0.0003 | 0.0188 | 0.0487 | 0.996 | 0.818 | 0.717 | 0.869 |
| 4-11-6              | 0.0043 | 0.0103 | 0.0631 | 0.964 | 0.920 | 0.597 | 0.858 |
| 4-12-6              | 0.0085 | 0.0429 | 0.0456 | 0.927 | 0.801 | 0.705 | 0.825 |
| 4-13-6              | 0.0000 | 0.0215 | 0.0109 | 0.999 | 0.879 | 0.931 | 0.944 |
| 4-14-6              | 0.0000 | 0.0631 | 0.0351 | 0.999 | 0.535 | 0.772 | 0.830 |

Figure 7. a) Best validation performance; b) training state; c) error histogram
solid waste prediction reported in the literature, the equations obtained were: (Singh & Satija, 2016), (Younes et al., 2015). The both models obtained in this study can be suitable models since they have higher values of the coefficient of determination for SW fractions generation prediction. In this study the most suitable is 4-13-6 model which have $R^2 = 0.94$, in others the values obtained were: $R^2 = 0.97$ (Younes et al., 2015), $R^2 = 0.82$ (Singh & Satija, 2016), $R^2 = 0.81$ (Zade & Noori, 2008) and $R^2 = 0.64$ (Shahabi et al., 2012).

### 2.3. Connection weights

The relative importance of each input variable in the network was determined by applying Garson’s algorithm (GA), according to Olden and Jackson (2002), Gevrey et al. (2003), Olden et al. (2004), Ibrahim (2013). Connection weights are presented in Table 4, and were obtained at the end of training process by using NIO for 4-11-6 model structure. In order to obtain product $P_{ij}$ (Table 5), the absolute value of the hidden-output layer connection weight was multiplied by the absolute value of the hidden input layer connection weight. $Q_{ij}$ values were obtained by dividing the $P_{ij}$ by the sum for all the input variables.

$Q_{ij}$ values were obtained by dividing the $P_{ij}$ by the sum for all the input variables. $S_i$ is the sum of $Q_{ij}$ for example

$$S_i = Q_{i1} + Q_{i2} + \ldots + Q_{i11} = 0.3254 + 0.2996 + \ldots + 0.0453 = 2.3744.$$  

The relative importance ($RI$) or distribution of all output weights attributable to the given input variable is obtained...
when $S_i$ is divided by the sum for all the input variables (Table 5). For example, $RI$ for the input neuron 2 is calculated with: $RI_2 = \frac{(S_2 \times 100)}{(S_1 + S_2 + S_3 + S_4)} = 30.5\%$.

Table 4. Input – hidden connection weights and hidden output connection weights for 4-11-6 model structure when NIO was applied

| Hidden neuron | Weights | Predictor variable | Output |
|---------------|---------|--------------------|--------|
|               |         | $x_1$ | $x_2$ | $x_3$ | $x_4$ |       |        |
| H1            | -1.4026 | 0.9104 | -1.2432 | -0.7540 | 0.8388 |
| H2            | 1.4299  | -1.8965 | 0.2355 | -1.2104 | 0.0700 |
| H3            | 0.5298  | -1.1203 | 1.6922 | 1.1345 | -0.0991 |
| H4            | -0.8609 | -1.5948 | 1.5247 | -0.6115 | -0.3684 |
| H5            | 1.7264  | -1.2900 | 0.0768 | -1.0994 | -0.6092 |
| H6            | -1.5473 | 1.2838 | 0.9669 | -0.6501 | 0.1077 |
| H7            | 0.7776  | 1.4473 | -1.0695 | 1.2690 | 0.5372 |
| H8            | 1.2637  | -0.9676 | 1.2375 | 1.3617 | 1.5464 |
| H9            | 0.3157  | 0.9004 | 2.2719 | 0.1028 | -1.7281 |
| H10           | 0.4427  | -0.7727 | -1.0438 | -1.4699 | -0.8554 |
| H11           | -0.1678 | -2.1348 | -0.9848 | -0.4150 | -0.7528 |

Table 4. Input – hidden connection weights and hidden output connection weights for 4-11-6 model structure when NIO was applied

| Hidden neuron | Connection weight Product $P_{ij}$ | $Q_{ij}$ |
|---------------|-----------------------------------|---------|
|               | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
| H1            | 1.1766 | 0.7637 | 1.0429 | 0.6325 | 0.3254 | 0.2112 | 0.2884 | 0.1749 |
| H2            | 0.1001 | 0.1327 | 0.0164 | 0.0847 | 0.2996 | 0.3973 | 0.0493 | 0.2536 |
| H3            | 0.0525 | 0.1111 | 0.1678 | 0.1125 | 0.1183 | 0.2502 | 0.3779 | 0.2534 |
| H4            | 0.3172 | 0.5875 | 0.5617 | 0.2253 | 0.1874 | 0.3473 | 0.3320 | 0.1331 |
| H5            | 1.0519 | 0.7860 | 0.0467 | 0.6699 | 0.4117 | 0.3076 | 0.0183 | 0.2622 |
| H6            | 0.1667 | 0.1383 | 0.1042 | 0.0700 | 0.3478 | 0.2886 | 0.2173 | 0.1461 |
| H7            | 0.4177 | 0.7775 | 0.5745 | 0.6817 | 0.1704 | 0.3171 | 0.2343 | 0.2780 |
| H8            | 1.9542 | 1.4964 | 1.9137 | 2.1058 | 0.2616 | 0.2003 | 0.2561 | 0.2818 |
| H9            | 0.5457 | 1.5562 | 3.9263 | 0.1777 | 0.0879 | 0.2507 | 0.6326 | 0.0286 |
| H10           | 0.3787 | 0.6610 | 0.8929 | 1.2574 | 0.1187 | 0.2072 | 0.2799 | 0.3941 |
| H11           | 0.1263 | 1.6070 | 0.7414 | 0.3124 | 0.0453 | 0.5765 | 0.2659 | 0.1121 |

| $S_i$ | 2.3744 | 3.3544 | 2.9526 | 2.3184 |
|-------|--------|--------|--------|--------|
| Relative importance ($RI_i$), % | 21.58 | 30.50 | 26.84 | 21.08 |
| Rank | 3 | 1 | 2 | 4 |
2016 (Ghinea et al., 2016). The study provides information for decision-makers to improve the management system of municipal solid waste and for future research. In future studies, we will also consider other input data, such as: economic data, waste collection system data and also, data on the behavior of residents. Since, in this study, we have performed the waste composition forecasting by applying ANN, in the following studies the results obtained in this paper will be compared with those obtained by applying regression analysis (Ghinea et al., 2016) and autoregressive integrated moving average (ARIMA).

Conclusions
In sustainable solid waste management, the solid waste prediction is a key aspect. By knowing the amount of SW fractions that will be generated the decision makers will be able to perform adequately planning and management of waste. In this paper ANN was used to investigate and predict solid waste composition in Iasi, Romania. The following conclusions can be drawn:
- NNNTS tool is a complex instrument which can be easily applied for SW fractions prediction;
- the results obtained by applying NARX and NIO models showed that both can be used for prediction, but NARX is more accurate;
- solid waste prediction depends on the variables which significantly influence the results. The proper input variables (like: number of residents, population aged 15 to 59 years, urban life expectancy, total municipal solid waste) should be selected considering the impact of input variables on model performance. Relevance, computational effort, training difficulty, dimensionality and comprehensibility are the main key considerations. The input variables must describe the behavior of the output variables and should have a minimum degree of redundancy. It is also important that the optimum input variable set does not contain uninformative variables;
- the optimum algorithm structure should be determined by choosing the smallest values RMSE for training, testing and validation (compared with the others models investigated) and the highest $R^2$ (0.94) and was a feedforward NARX network (4 inputs, 13 neurons and 6 outputs) with Levenberg-Marquardt training algorithm;
- the NARX model can be used and applied easily also in other cities of the country because waste composition is similar (the percentages are the same), instead the waste amount produced is different and but is closely related to the inputs variables, which may slightly vary from one region to another.
- with the ANN methodology can be determined the linear regression between the inputs and targets data and also a time series response.

Acknowledgements
This work was elaborated with the support of a Grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-III-P4-ID-PCE-2016-0683, contract 65/2017.

Funding
This work was supported by the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, under Grant [number PN-III-P4-ID-PCE-2016-0683].

Author contributions
CG conceived the study and is responsible for the design and development of the data analysis. CG and PC were responsible for data collection and data interpretation. CG wrote the draft of the article. MG revised the data processing and paper, and give suggestions and comments of results and explanation for expanded analysis of results.

Disclosure statement
We declare that we did not have any competing financial support, professional, or personal interests from other parties.

References
Abbas, M., & El Hanandeh, A. (2016). Forecasting municipal solid waste generation using artificial intelligence modelling approaches. Waste Management, 56, 13–22. https://doi.org/10.1016/j.wasman.2016.05.018
Abbas, M., Abduli, M.A., Omidvar, B., & Baghvand, A. (2014). Results uncertainty of support vector machine and hybrid of wavelet transform-support vector machine models for solid waste generation forecasting. Environmental Progress & Sustainable Energy, 33(1), 220–228. https://doi.org/10.1002/ep.11747
Adamović, V. M., Antanasijević, D. Z., Ristić, M. D., Perić-Grujić, A. A., & Pocajt, V. V. (2016). Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis. Environmental Science and Pollution Research, 24(1), 299–311. https://doi.org/10.1007/s11356-016-7767-x
Agatonović-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis, 22(5), 717–727. https://doi.org/10.1016/S0731-7085(99)00272-1
Ali, S. A., & Ahmad, A. (2019). Forecasting MSW generation using artificial neural network time series model: a study from metropolitan city. SN Applied Sciences, 1, 1338. https://doi.org/10.1007/s42452-019-1382-7
Antanasijević, D., Pocajt, V., Popović, I., Redžić, N., & Ristić, M. (2013). The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. Sustainability Science, 8(1), 37–46. https://doi.org/10.1007/s11625-012-0161-9
Batinic, B., Vukmirovic, S., Vujic, G., Stanislavljevic, N., Ubarvin, D., & Vukmirovic, G. (2011). Using ANN model to deter-
mine future waste characteristics in order to achieve specific waste management targets – case study of Serbia. *Journal of Scientific & Industrial Research*, 70(07), 513–518.

Boussaada, Z., Cureka, O., Remaci, A., Cambliong, H., & Beliaoj, N. M. (2018). A Nonlinear Autoregressive Exogenous (NARX) neural network model for the prediction of the daily direct solar radiation. *Energies*, 11(3), 620. https://doi.org/10.3390/en11030620

Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, 160(3), 249–264. https://doi.org/10.1016/S0304-3800(02)00257-0

Ghinea, C. (2012). Waste management models and their application to sustainable management of recyclable waste [PhD Thesis]. Gheorghe Asachi Technical University of Iasi, Romania.

Ghinea, C., & Gavrilescu, M. (2010). Decision support models for solid waste management – an overview. *Environmental Engineering and Management Journal*, 9(6), 869–880. https://doi.org/10.30638/eemj.2010.116

Ghinea, C., & Gavrilescu, M. (2016). Costs analysis of municipal solid waste management scenarios: IASI – Romania case study. *Journal of Environmental Engineering and Landscape Management*, 2(4), 185–199. https://doi.org/10.3846/16486897.2016.1173041

Ghinea, C., & Gavrilescu, M. (2019). Solid waste management for circular economy. Challenges and opportunities in Romania – The case study of Iasi County. In M. L. Franco-Garcia, J. Carpio-Aguilar, & H. Bressers (Eds.), *Greening of industry networks studies: Vol. 6. Towards zero waste* (pp. 25–60). Springer. https://doi.org/10.1007/978-3-319-92931-6_3

Ghinea, C., Dragoi, E. N., Comanita, E.-D., Gavrilescu, M., Campean, T., Curteanu, S., & Gavrilescu, M. (2016). Forecasting municipal solid waste generation using prognostic tools and regression analysis. *Journal of Environmental Management*, 182, 80–93. https://doi.org/10.1016/j.jenvman.2016.07.026

Ghinea, C., Petraru, M., Bressers, J. T. A., & Gavrilescu, M. (2012). Environmental evaluation of waste management scenarios – significance of the boundaries. *Journal of Environmental Engineering and Landscape Management*, 20(1), 76–85. https://doi.org/10.3846/16486897.2011.644665

Goel, S., Ranjan, V. P., Bardhan, B., & Hazra, T. (2017). Forecasting solid waste generation rates. In D. Sengupta & S. Agrahari (Eds.), *Modelling trends in solid and hazardous waste management* (pp. 35–64). Springer, Singapore. https://doi.org/10.1007/978-981-10-2410-8_3

Guthrie, D. (2008). *Unsupervised detection of anomalous text*. University of Sheffield. https://citeserv.ist.psu.edu/viewdoc/download?doi=10.1.1.396.6322&rep=rep1&type=pdf

Han, S. H., Kim, K. W., Kim, S. Y., & Youn, Y. K. (2018). Artificial neural network: Understanding the basic concepts with understanding mathematics. *Dementia and Neurocognitive Disorder*, 17(3), 83–89. https://doi.org/10.12779/dnd.2018.17.3.83

Henriques, R. S., & Coelho, L. M. G. (2019). A multi-objective optimization model to support the municipal solid waste management. *Environmental Engineering and Management Journal*, 18(5), 1077–1088. https://doi.org/10.30638/eemj.2019.104

Ibrahim, O. M. (2013). A comparison of methods for assessing the relative importance of input variables in artificial neural networks. *Journal of Applied Sciences Research*, 9(11), 5692–5700. http://www.aensilweb.com/old/jasr/jasr/2013/5692-5700.pdf

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

Kumar, J. S., Subbaiah, K. V., & Prasada Rao, P. V. V. (2011). Prediction of municipal solid waste with RBF Net Work – A case study of Eluru, A. P. India. *International Journal of Innovation, Management and Technology*, 2(3), 238–243.

Kumar, A., & Samadder, S. R. (2017). An empirical model for prediction of household solid waste generation rate – a case study of Dhanbad, India. *Waste Management*, 68, 3–15. https://doi.org/10.1016/j.wasman.2017.07.034

Li, Y. P., Sun, Y., Huang, G. H., & Zhao, F. (2016). An interval-parameter queuing model for planning municipal solid waste management system with cost-effective objective. *Environmental Engineering and Management Journal*, 15(8), 1673–1687. https://doi.org/10.30638/eemj.2016.180

Marcjasz, G., Uniejewski, B., & Weron, R. (2019). Probabilistic electricity price forecasting with NARX networks: Combine point or probabilistic forecasts? *International Journal of Forecasting*, 36(2), 466–479. https://doi.org/10.1016/j.ijforecast.2019.07.002

MathWorks. (2016). *Neural Network Toolbox* User’s Guide. https://www.mathworks.com/help/pdf_doc/nnet/nnet Ug.pdf

Menezes, J. M. P. Jr., & Barreto, G. A. (2008). Long-term time series prediction with the NARX network: An empirical evaluation. *Neurocomputing*, 71(16–18), 3335–3343. https://doi.org/10.1016/j.neucom.2008.01.030

National Institute of Statistics. (2015). Iasi county statistics, county statistics. NIS. http://www.iasi.insse.ro/main.php?id=373/NIST/SEMETECH. (2015). *e-Handbook of statistical methods*. National Institutes of Standards and Technologies (NIST), US. http://www.itl.nist.gov/div898/handbook/

Noori, R., Abdoli, M., Jalili Ghazizade, M., & Samieifard, R. (2009). Comparison of Neural Network and Principal Component-Regression Analysis to predict the solid waste generation in Tehran. *Iranian Journal of Public Health*, 38(1), 74–84.

Noori, R., Karbassi, A., & Salman Sabahi, M. (2010a). Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management*, 91(3), 767–771. https://doi.org/10.1016/j.jenvman.2009.10.007

Noori, R., Khakpour, A., Omidvar, B., & Farokhnia, A. (2010b). Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications*, 37(8), 5856–5862. https://doi.org/10.1016/j.eswa.2010.02.020

Noori, R., Karbassi, A. R., Mehdizadeh, M., & Samieifard, R. (2018). Artificial neural networks studies: Vol. 6. Towards zero waste. In D. Sengupta & S. Agrahari (Eds.), *Modelling trends in solid and hazardous waste management* (pp. 25–60). Springer, Singapore. https://doi.org/10.1007/978-981-10-2410-8_3

Olden, J. D., & Jackson, D. A. (2002). Illuminating the “black box”: A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154(1–2), 135–150. https://doi.org/10.1016/S0304-3800(02)00064-9

Olden, J. D., Joy, M. K., & Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, 178(3–4), 389–397. https://doi.org/10.1016/j.ecolmodel.2004.03.013
