Journal of Material Cycles and Waste Management (2018) 20:1736–1750
https://doi.org/10.1007/s10163-018-0741-6

ORIGINAL ARTICLE

An optimized artificial neural network model for the prediction of rate of hazardous chemical and healthcare waste generation at the national level

Vladimir M. Adamović[1] · Davor Z. Antanasijević[2] · Mirjana Đ. Ristić[3] · Aleksandra A. Perić-Grujić[3] · Viktor V. Pocajt[3]

Received: 26 May 2017 / Accepted: 24 April 2018 / Published online: 30 April 2018

Abstract
This paper presents a development of general regression neural network (a form of artificial neural network) models for the prediction of annual quantities of hazardous chemical and healthcare waste at the national level. Hazardous waste is being generated from many different sources and therefore it is not possible to conduct accurate predictions of the total amount of hazardous waste using traditional methodologies. Since they represent about 40% of the total hazardous waste in the European Union, chemical and healthcare waste were specifically selected for this research. Broadly available social, economic, industrial and sustainability indicators were used as input variables and the optimal sets were selected using correlation analysis and sensitivity analysis. The obtained values of coefficients of determination for the final models were 0.999 for the prediction of chemical hazardous waste and 0.975 for the prediction of healthcare and biological hazardous waste. The predicting capabilities of the models for both types of waste are high, since there were no predictions with errors greater than 25%. Also, results of this research demonstrate that the human development index can replace gross domestic product and in this context even represent a better indicator of socio-economic conditions at the national level.

Keywords Hazardous waste · Chemical waste · Healthcare waste · Medical waste · Artificial neural networks

Introduction

A waste can be classified as hazardous when it possesses one or more of the following characteristics: explosive, oxidizing, flammable, irritant, toxic, carcinogenic, corrosive, infectious, toxic for reproduction, mutagenic, releases toxic gas in contact with water, air or an acid, sensitizing, eco-toxic and reactive [1, 2]. Hazardous waste comes from all economic sectors, from common households to the outputs from waste treatment plants (secondary waste), but excludes radioactive waste, decommissioned explosives, waste waters, or animal by-products (except those destined for incineration, landfilling, or use in biogas or composting plants) [3]. Its quantity is relatively small compared to the generated total waste, but it is potentially very damaging to both the environment and human health. For this reason, hazardous waste is subject to stricter legislations and controls [4], since its poor management can lead to a significant degradation of soil, groundwater, human and animal health and the environment in general [5].

Generation of hazardous waste belongs to the Sustainable Development Indicators (SDI) set, as it has been chosen for the assessment of the progress towards the objectives and targets of the EU Sustainable Development Strategy [6]. Due to harm it may cause, hazardous waste is a very sensitive topic and a particularly important field within the domain of environment statistics [5, 7].

Total hazardous waste covers chemical and healthcare (or medical) wastes, mineral and solidified wastes (mineral waste from construction and demolition, combustion waste, soils, dredging spoils, mineral wastes from waste treatment and stabilized wastes and other mineral wastes), recyclable wastes (metal waste, ferrous, paper and

✉ Vladimir M. Adamović
v.adamovic@itnms.ac.rs

[1] Institute for Technology of Nuclear and other Mineral Raw Materials, Bulevar Franš d'Eperea 86, Belgrade 11000, Serbia

[2] Innovation Center of the Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, Belgrade 11120, Serbia

[3] Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, Belgrade 11120, Serbia

cardboard waste, rubber, plastic, etc.), equipment (waste-containing PCB, discarded equipment, discarded vehicles, batteries and accumulators wastes, etc.) and a mixture of ordinary wastes (household and similar wastes, mixed and undifferentiated materials, common sludge, etc.). Directly after mineral and solidified waste, chemical and healthcare wastes (subtotal) represent the most significant category of hazardous waste in the European Union (about 40%) [8].

The main types of hazardous chemical wastes are spent solvents, acid, alkaline or saline wastes, used oils, chemical wastes in the strictest sense (chemical deposits and residues and chemical preparation wastes), industrial effluent sludge, and sludge and liquid wastes from waste treatment. Even though only a small fraction of more than 80,000 chemicals registered by the U.S. Environmental Protection Agency (EPA) have been tested to evaluate their potential risk to human health, wildlife and the environment [9], their potentially harmful impact can be reflected in the fact that among those tested, 216 chemicals were identified to have been associated with increases in mammary gland tumours in at least one study [10].

Except general waste (municipal solid waste), all wastes generated in any healthcare establishment are defined as healthcare wastes, but not all healthcare wastes are characterized as hazardous wastes. There are several different definitions of hazardous healthcare waste, but by the most commonly used one [11] defines hazardous healthcare waste as waste which is generated during the diagnosis, treatment, or immunization of humans or animals, in research related to these activities and in production or testing of biological resources, which may be contaminated with infectious organisms, or could lead to injury or other health impact. Hazardous healthcare waste includes used bandages, laboratory waste glass, used surgical gloves and instruments (scalpels, etc.), used needles and syringes, microbiological agents and culture, amputated organs, blood and blood products, chemicals, pharmaceuticals, animal carcasses, pathological waste, etc. [12, 13].

This waste should be segregated at the point of generation between hazardous healthcare waste and municipal solid waste (MSW). The World Health Organisation (WHO) estimates that only between 10% and 25% of total healthcare waste is potentially hazardous waste, and a potential source of infection, injury or other health impact [14–16].

The ability to accurately predict the amount of chemical and medical waste is extremely important for planning and development of strategies for hazardous waste management at the local and national levels [7, 17], which corresponds to the aim of this study—the development of models for the prediction of annual generation of hazardous chemical waste and hazardous healthcare waste at the national level. Since these two types of waste are different in the manner and origin of generation, separate models were developed for each of them.

General Regression Neural Network (GRNN), a type of Artificial Neural Network (ANN), was used for the development of the both models. ANNs are computerized mathematical tools inspired by the behaviour of human neurons and the electrical signals that they convey between input, processing of information and subsequent output from the brain. Learning from experience (or examples) is crucial for ANNs [18].

This survey is a continuation of the previous research in which the amounts of MSW generated at the national level were investigated [19, 20].

Materials and methods

Available data

Output data and methodology

Obtaining reliable data on annual quantities of hazardous waste at national level represented a starting point for the modelling of hazardous waste generation. For this purpose, data from the Eurostat database were used [8]. The biannual quantities of generated chemical hazardous waste (CHW) and healthcare and biological hazardous waste (HCBHW) at the national level in kg per capita were used separately as single output variables.

The data relating to the generation of waste are collected from the EU Member States on the basis of the Regulation on waste statistics (EC) No. 2150/2002, amended by the Commission Regulation (EU) No. 849/2010 [8]. The data collection methods are surveys, administrative sources, statistical estimations or varying combination of these methods. The Member States are free to decide which method will be used. The database covers the European Union (28 countries), Norway as a member of the European Economic Area and Turkey as an EU candidate country.

Overall, 3.7% of the total waste in the EU-28 in 2014 was classified as hazardous waste and in most of the European countries it varies between 1 and 8% of total waste. Exceptions are Estonia (47.7% in 2014), and Norway (12.1%). In Estonia, the high share of hazardous waste is due to energy production from oil shale [8], while in Norway the higher percentage can be attributed to oil drilling activities on the shelf and the manufacturing industry [21].

Unlike municipal solid waste (MSW), for which national annual data have been collected for more than 20 years, national biannual data for CHW and HCBHW generation are only available from the year 2004. During that period, in some of the countries methodology for data collection was changed, which resulted in huge differences among the data reported by the same country in different years. These differences can be an indicator of poor data collection or improper waste management practices in the respective country [22] or even be attributed to the differences in how HCBHW and CHW are defined and categorized in different countries [11] leading to more profound differences in the reported amounts of generated waste [22].

Due to the large differences between the data for the output variables (generated CHW or generated HCBHW) for the same country in different years, it was obvious that the data for some countries had to be treated as outliers. Such outliers can lead to misspecification, biased parameter estimation, incorrect results and poor overall performance of the model [23]. Hence, before modelling and analysis, it is important to identify the outliers and remove them from the dataset [24].

Among the data of CHW generation, outliers were observed in the following countries: Belgium, Bulgaria, Ireland, Cyprus, Latvia and Romania. As an example of inconsistency and an outlier, in Cyprus the largest value of CHW generation of 114.54 kg per capita was recorded in 2004, but the average in other years was 8.59 kg per capita.

In case of the HCBHW generation, there are even more outliers in the datasets, namely in the case of Belgium, Ireland, Croatia, Cyprus, Latvia, Luxembourg, Romania and Slovenia. The most evident cases are Belgium, where 48.65 kg of generated HCBHW per capita was recorded in 2012, while the average in other years was 1.95 kg per capita, and Croatia, where the recorded amount of generated HCBHW in 2006 was 114.35 kg per capita, but the average for other years was 0.71 kg per capita.

Also, there are two countries with reported amount of 0 tonne of the HCBHW per year: Norway (for all observed years, except 2014) and Turkey (for 2006). The data from these countries have not been used for the modelling of HCBHW.

Due to lacks of some input data, Malta has not been taken into account for the modelling of the CHW generation.

After the data for the previously mentioned countries was removed, the data from the remaining 23 countries (Table 1) and 19 countries (Table 2) were used for the modelling of CHW and HCBHW generation, respectively.

The data for 5 years (2004, 2006, 2008, 2010 and 2012) were used for the training and validation of the models, and the data for 2014 were used to test the prediction capability of the models.

Table 1 Descriptive statistic of the biannual CHW generation from 2004 to 2014

Country	Chemical hazardous waste (CHW) (kg per capita)			
	Mean	Stand. dev.	Minimum	Maximum
Czech Republic	58.82	3.10	53.67	62.54
Denmark	32.40	9.12	23.25	45.26
Germany	83.38	815	76.89	96.97
Estonia	1085.11	154.14	889.70	1261.29
Greece	10.34	4.02	6.52	17.51
Spain	43,26	7.97	33.97	54.09
France	46.84	1.97	43.18	48.64
Croatia	11.56	5.90	5.52	21.90
Italy	60.91	7.53	48.08	71.74
Lithuania	18.37	3.58	14.82	23.46
Luxembourg	67.08	6.71	61.28	80.06
Hungary	35.44	15.55	26.20	66.98
Netherlands	84.37	12.05	65.47	96.57
Austria	46.98	3.54	41.19	50.68
Poland	26.19	11.17	18.07	46.78
Portugal	91.16	145.45	23.97	387.14
Slovenia	31,26	5.18	26.24	38.93
Slovakia	38.67	3.68	33.56	42.90
Finland	96,72	12.00	78.72	113.98
Sweden	73.63	6.89	64.32	85.46
United Kingdom	44.31	12.11	28.42	61.01
Norway	133.71	48.22	83.07	189.76
Turkey	7.61	3.85	0.15	1024
The countries not included in the CHW model				
Belgium	137.46	39.25	91.12	186.85
Bulgaria	47.17	32.43	6.84	86.45
Ireland	148.80	141.24	39.53	334.39
Cyprus	26.25	43.37	5.35	114.54
Latvia	16.40	6.69	5.76	25.20
Malta	64.53	33.34	27.53	107.38
Romania	11.26	4.29	6.10	15.16

Input data

The lists of input variables selected for the creation of CHW and HCBHW models are presented in Table 3. The data were obtained from: European Statistical Office—Eurostat, World Bank, United Nations Development Program (UNDP), Organisation for Economic Co-operation and Development (OECD) and from several national databases. Detailed description of each input is given in the Appendix A.

Unlike MSW generation, for which there is a significant number of forecasting models [25], scientific literature relating to the indicators and modelling of hazardous waste generation is very scarce [17].

There are just a few papers dealing with modelling of hazardous medical waste generation [26–30] in which the following variables were used as inputs: number of hospitals, number of occupied beds in hospitals, number of beds per capita, number of inpatients, number of total patients, population growth, visits at hospitals and other medical institutions, number of children, and average

Table 2 Descriptive statistic of the biannual HCBHW generation from 2004 to 2014

Country	Healthcare and biological hazardous waste (HCBHW) (kg per capita)			
	Mean	Stand. dev.	Minimum	Maximum
Bulgaria	0.21	0.07	0.12	0.28
Czech Republic	2.29	0.37	1.82	2.62
Denmark	1.21	0.59	0.79	2.38
Germany	0.12	0.02	0.08	0.15
Estonia	0.25	0.16	0.03	0.46
Greece	1.44	0.11	1.33	1.57
Spain	2.68	1.00	1.74	4.26
France	3.94	2.74	1.34	6.68
Italy	2.38	0.11	2.21	2.54
Lithuania	0.27	0.14	0.08	0.43
Hungary	2.95	2.07	1.32	7.08
Malta	0.80	0.14	0.56	0.91
Netherlands	0.58	0.10	0.39	0.67
Austria	0.86	0.90	0.15	2.45
Poland	0.96	0.32	0.48	1.31
Portugal	5.09	6.80	1.25	18.93
Slovakia	5.93	3.00	2.34	9.84
Sweden	1.16	1.72	0.40	4.67
United Kingdom	4.68	1.01	3.50	5.93
The countries not included in the HCBHW model				
Belgium	9.73	19.07	1.39	48.65
Ireland	1.40	1.94	0.01	4.12
Croatia	19.65	46.39	0.36	114.35
Cyprus	0.77	0.37	0.42	1.29
Latvia	0.74	0.81	0.14	2.31
Luxembourg	2.47	2.91	0.65	7.79
Romania	0.64	0.12	0.52	0.82
Slovenia	3.20	2.34	0.46	6.30
Finland	0.96	0.53	0.01	1.64
Norway	0.01	0.03	0	0.07
Turkey	0.60	0.41	0	1.11

life expectancy. Papers dealing with prediction of hazardous chemical waste are even rarer. Indicators used in these studies are amount of the crude oil processed, gross domestic product, urban population, energy consumption [17, 31].

In this study for development of both models the following universal indicators were used: gross domestic product (GDP) as an economic indicator, domestic material consumption (DMC) for technical, human development index (HDI) for socio-economic and urban population (UP) as a demographic indicator. In addition to these common indicators, several specific indicators were used as well:

- For the CHW prediction model:
- Industrial indicators: value added of industry (VAI) and final energy consumption in chemical and petrochemical industry (FECCPI);
- Agricultural indicators: agriculture land (AL); value added of agriculture (VAA) and fertilizer consumption (FC), and
- Scientific indicators: scientist and engineers (SE) and total intramural research and development expenditure (TIRDE).
- For the HCBHW prediction model:
- Medical indicators: available beds in hospitals (ABH); health expenditure (HE); immunization, measles (IM); incidence of tuberculosis (IT); people having a long-standing illness or health problem (PHLSIHP); bed-days in hospitals (BDH) and number of doctors (ND);
- Demographic indicators: population of the ages 65 and above (PA65+); victims in road accidents (VRA); life expectancy at birth (LEB) and death rate, crude (DRC) and
- Socio-economic indicator: inability to face unexpected financial expenses (IFUFE).

Correlation analysis (CA)

Correlated data can reduce the distinctiveness of data representation and it can lead to confusion in the ANN modelling during the learning process [32].

Considering that the presence of correlated data is not advisable, it is necessary to determine which data are correlated and to remove a proportion of it. Correlation among two variables could be easily determined using Pearson's correlation analysis:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y})}{(n-1)\cdot\sigma_x\cdot\sigma_y}, \tag{1}$$

where r is the correlation coefficient, x and y are the variables whose correlation is being tested, $\bar{x}$ and $\bar{y}$ are the mean values of the tested population ($i = 1,2\ldots n$) and σ_x and σ_y are the standard deviations of x and y variables, respectively [33].

In a situation where two input variables are highly correlated, i.e. if their mutual correlation coefficients are higher than 0.8 [34], one of them can be removed without negatively affecting the ANN performance [35].

General regression neural network (GRNN)

GRNN is a type of artificial neural networks (ANNs) originally proposed by Specht [18]. The main advantages

Table 3 The list of input variables selected for the creation of CHW and HCBHW models

Input variable	Unit	CHW	HCBHW
Gross domestic product (GDP) at market prices	Purchasing power standard per capita	X	X
Domestic material consumption (DMC)	*t* per capita	X	X
Human development index (HDI)	–	X	X
Urban population (UP)	% of total	X	X
Value added of industry (VAI)	Share of total	X	–
Final energy consumption in chemical and petrochemical industry (FECCPI)	kg of oil equivalent per capita	X	–
Agriculture land (AL)	% of land area	X	–
Value added of agriculture (VAA)	Share of total	X	–
Fertilizer consumption (FC)	kg per hectare of arable land	X	–
Scientists and engineers (SE)	% of total population	X	–
Total intramural research and development expenditure (TIRDE)	Euro per inhabitant	X	–
Available beds in hospitals (ABH)	Per 100,000 people	–	X
Health expenditure (HE)	% of GDP	–	X
Population of the ages 65 and above (PA65+)	% of total	–	X
Victims in road accidents (killed + injured) (VRA)	Per million inhabitants	–	X
Life expectancy at birth (LEB)	Total years	–	X
Inability to face unexpected financial expenses (IFUFE)	% of total population	–	X
Immunization, measles (IM)	% of children ages 12–23 months	–	X
Incidence of tuberculosis (IT)	Per 100,000 people	–	X
Death rate, crude (DRC)	Per 1,000 people	–	X
People having a long-standing illness or health problem (PHLSIHP)	% of total	–	X
Bed days in hospitals (BDH)	Per capita	–	X
Number of doctors (ND)	Per 100,000 inhabitants	–	X

of GRNN are its simplicity, its non-parametric nature and its ability to make accurate predictions even with small or often incomplete data series [36]. GRNN is the so-called "black-box" tool (Fig. 1) which means that the relation between input and output data is not necessarily known [37].

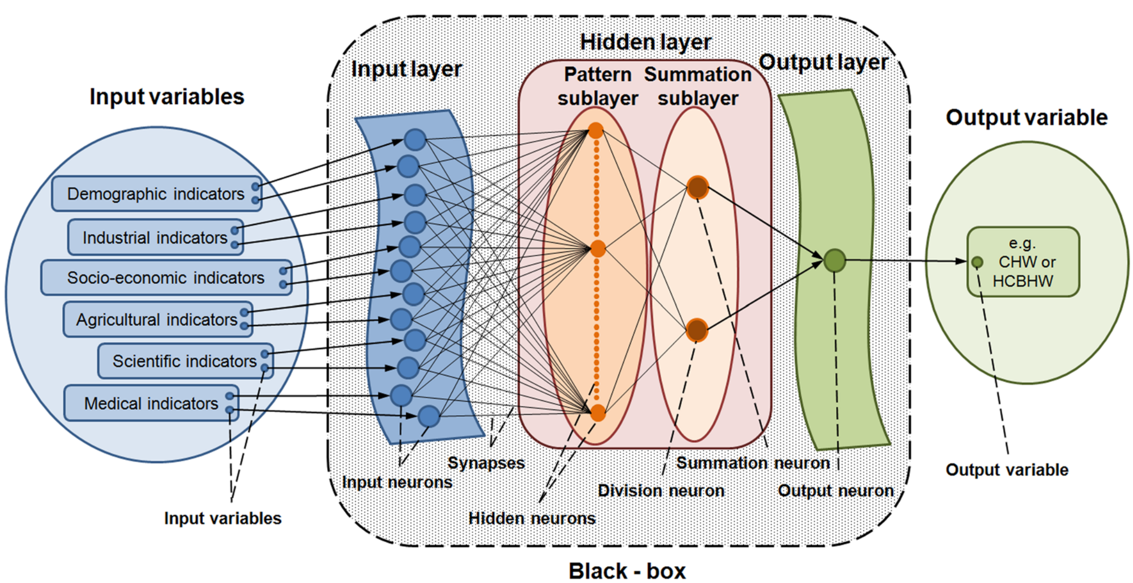


Fig. 1 GRNN architecture

GRNN is mainly used to solve nonlinear problems which are based on the estimation of a probability distribution function [38]. This is a computationally very efficient approach since it is a one-pass supervised machine learning algorithm [39, 40]. Specifically, GRNN was used in the modelling of many different environment phenomena because of its advantages over other forecasting methods [19, 41–44].

GRNN is actually a type of probabilistic neural network (PNN) [45]. However, while PNN uses a probability density function to predict the discrete values of the probability of certain events, with a GRNN this principle has been extended and applied to determine the values of continuous variables. In other words, a PNN performs classification where the target variable is categorical, whereas GRNN performs regression where the target variable is continuous [46].

The GRNN architecture is composed of four layers: input, pattern, summation and output layers (Fig. 1). Each input variable is presented by one neuron in the input layer and each pattern in the training dataset is presented by one neuron in the pattern layer. The output layer has one neuron for each output variable, and the summation layer has one neuron more than there are in the output layer. Usually, there is one neuron in the output layer and two (the summation and division neurons) in the summation layer. Also, it could be said that the pattern layer and the summation layer are parts of the hidden layer.

The network's output (y) can be derived using Eq. 2:

$$y = \frac{S_1}{S_2}. \tag{2}$$

S_1 represents the sum of the weighted outputs of the pattern layer, computed by the summation neuron:

$$S_1 = \sum_{j=1}^{k} y_j \cdot f(D_j). \tag{3}$$

S_2 are the un-weighted outputs of the pattern layer, computed by the division neuron:

$$S_2 = \sum_{j=1}^{k} f(D_j), \tag{4}$$

where k is the number of data patterns, y_j represents the measured values of the output variables and $f(D_j)$ is the activation function (Eq. 5):

$$f(D_j) = \exp\left(\frac{-D_j}{2 \cdot \sigma_f^2}\right). \tag{5}$$

In Eq. 5, σ_f is the smoothing factor, while D_j shows the distance between the training patterns and it can be calculated using the Euclidean distance method:

$$D_j = \sqrt{\sum_{i=1}^{n} (w_{ij} - x_i)^2}, \tag{6}$$

where n is the number of inputs, w_{ij} represents the weight of the neurons and x_i is the input data pattern.

The smoothing factor (σ_f) determines the accuracy of the GRNN, and in general, if it is closer to 0, the regression surfaces are smoother [43]. The smoothing factor is the only parameter that is unknown in the GRNN algorithm, and it is determined within the process of the training of the network.

The links between the neurons in the input layer and the neurons in the pattern layer are made using a linear scaling function, while for the links between the neurons in pattern layer and the neurons in the summation layer an exponential activation function was used [18].

In the current case, this means that the value of GRNN weights between the first two layers was directly set by the input values presented in the training subset, while the weights between hidden layers were set by the CHW and HCBHW values from the training subset. The optimal smoothing factor was determined by applying generic algorithm to the data from the validation subset [47].

Model evaluation criteria

Various performance metrics have been proposed and used to quantify the accuracy of modelling results [42, 48, 49]. There is no set of universal performance metrics for the assessment of models, but usual practice involves the use of several different evaluation criteria [50]. In this study, the following evaluation criteria were applied: the coefficient of determination (R^2), the root mean squared error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the percentage of prediction within the factors of 1.1 and 1.25 (FA1.1 and FA1.25) (for details see Appendix B).

Results and discussion

Selection of input variables based on correlation analysis

The input data presented in the upper and middle parts of Table 3 were used for the modelling of the generation amount of chemical hazardous waste (CHW) in 23 European countries. The data from 2004 to 2012 were used for training and validation of the GRNN (CHW) model whilst the data from 2014 were reserved for testing the model. The data were randomly divided in the ratio of 4:1 (92 training and 23 validation patterns).

Correlation analysis results, for the CHW model dataset, indicate high mutual correlation, higher or close to ±0.8, between the human development index (HDI) on one side and the gross domestic product at market prices (GDP), value added of agriculture (VAA) and total intramural research and development expenditure (TIRDE) on the other. That could be expected, since HDI represents a composite statistic used to evaluate countries in terms of three main indicators of economic and social welfare: income, health and education attainments [51].

Between the human development index and gross domestic product per capita, HDI was chosen because it is a better indicator than GDP per capita for the measurement of national progress [51, 52]. Also, to reduce the number of input variables as much as possible, HDI was selected instead of three other previously mentioned indicators.

Therefore, for the creation of the CHW prediction model all input data from the upper and middle parts of Table 3, except gross domestic product (GDP), value added of agriculture (VAA) and total intramural research and development expenditure (TIRDE) were used. This model is denoted as CA-GRNN (CHW). It should be noted that during the development of the model various combinations with the abovementioned input parameters were tested and the best performance was achieved when HDI was included, and GDP, VAA and TIRDE were excluded from the input dataset.

For the modelling of the amount of healthcare and biological hazardous waste (HCBHW) in 19 European countries, the input data presented in the lower and middle parts of Table 3 were used. Again, the biannual data from 2004 to 2012 were used for training and validation of the GRNN (HCBHW) model and the data from 2014 for testing the model. The data were randomly divided in the proportion of 4:1 (76 training and 19 validation patterns).

High mutual correlations (above and close to 0.8) are noted among the human development index (HDI) on one side and gross domestic product at market prices (GDP), health expenditure (HE) and life expectancy at birth (LEB) on the other.

In this case, a priority was again given to HDI, and consequently, GDP, HE and LEB were removed from the dataset. The model is denoted as CA-GRNN (HCBHW). As in the previous case, various combinations of input parameters were tested and the best performance was achieved when HDI was used, and GDP, HE and LEB were omitted from the input dataset.

A linear scale function was used for the connection between the input and hidden neurons, while for linking of the hidden and summation neurons, an exponential activation function was used. The distance among training patterns was calculated using Euclidean distance method, whilst generic algorithm was used for the determination of the smoothing factor (Table 4).

Details on created CHW and HCBHW models are presented in Table 4.

Sensitivity analysis (SA) and model optimization

To optimize ANN model after correlation analysis, sensitivity analysis was performed to determine the effect of every single remaining input variable on the predicting capabilities of the model. Sensitivity analysis can be used to determine relative importance of input variables on the output, which allows the removal of variables that do not have a significant effect on the performance of the model. Thus, the network becomes more compact [53].

Sensitivity analysis can be performed based on individual smoothing factors (ISFs). The ISFs values are ranked by size

Table 4 Details on the created CHW and HCBHW models

Model		CHW		HCBHW	
		CA-GRNN	SA-CA-GRNN	CA-GRNN	SA-CA-GRNN
Datasets	Training (2004–2012)	92 patterns		76 patterns	
	Validation (2004–2012)	23 randomly selected patterns		19 randomly selected patterns	
	Test (2014)	23 patterns		19 patterns	
Input selection		Correlation analysis	Correlation and sensitivity analysis	Correlation analysis	Correlation and sensitivity analysis
Architecture parameters (number of neurons)	Input	8	6	13	10
	Summation	2			
	Output	1			
Training parameters	Scale function	Linear [0, 1]			
	Activation function	Exponential			
	Distance metric	Euclidean			
	Algorithm	Genetic			
	Termination	Automatic after 20 generations with no improvement of 1%			

from 0 to 3, wherein the higher the value of ISF, the higher the significance of the input [47].

For the second model for the prediction of CHW generation, only input variables with an ISF higher than 1 were used. Variables: domestic material consumption (DMC), with ISF = 0.62 and scientists and engineers (SE) with ISF = 0.08 have been removed from the input dataset. Input data that were used for the creation of the second CHW prediction model, denoted as SA-CA-GRNN (CHW), are: human development index (HDI), urban population (UP), value added in industry (VAI), final energy consumption in chemical and petrochemical industry (FECCPI), agricultural land (AL) and fertilizer consumption (FC).

Input variable with ISF higher and close to 1 were used for the second model for the prediction of HCBHW generation, denoted as SA-CA-GRNN (HCBHW). Variables: domestic material consumption DMC (ISF = 0.12), victims (killed and injured) in road accidents (VRA) (ISF = 0.06) and death rate, crude (DRC) (ISF = 0.05) have been removed from the input dataset, hence the following ten variables remained: human development index (HDI), urban population (UP), available beds in hospitals (ABH), population of the ages 65 and above (PA65+), inability to face unexpected financial expenses (IFUFE), immunization, measles (IM), incidence of tuberculosis (IT), people having a long-standing illness or health problem (PHLSIHP), bed-days in hospitals (BDH) and number of doctors per 100,000 inhabitants (ND).

Details on the created CHW and HCBHW models are presented in Table 4.

CHW and HCBHW model results

Performance metric values for the models created for prediction of amount of CHW and HCBHW generation on national level are presented in Table 5. As mentioned before, the designator CA-GRNN refers to the models where optimization of input parameters was performed using correlation analysis, while SA-CA-GRNN refers to the models obtained using both correlation and sensitivity analysis.

After correlation analysis was performed, various combinations of input parameters were tested to develop models for the prediction of CHW and HCBHW generation. In both cases, the best performance was achieved by the models developed using HDI as one of the input parameters and by dropping out GDP, VAA and TIRDE in the CHW modelling, as well as GDP, HE and LEB in the HCBHW modelling.

It was also of interest to compare performance indicators of the models developed using HDI versus the models developed using GDP as input parameter. When GDP was used, in the case of CHW modelling, R^2, IA and FA1.2 were equal to the corresponding indicators from modelling with HDI (model CA-GRNN (CHW)—Table 5), but the other indicators were not as good (RMSE = 20.49 kg pc; MAE = 10.31 kg pc; MAPE = 13.20%; FA1.1 = 47.83% and FA1.25 = 96.00%). In the case of HCBHW modelling, when GDP was used instead of HDI, the percentages of predictions within a factor of the observed values (FA1.1; FA1.2 and FA1.25) were equal to the corresponding indicators obtained by modelling with HDI (CA-GRNN (HCBHW) model - Table 5), but the other indicators were slightly lower (R^2 = 0.965; RMSE = 0.32 kg pc; MAE = 0.21 kg pc and MAPE = 13.48%).

Table 5 Performance indicators of created GRNN models for prediction of amount of CHW and HCBHW generation

	CA-GRNN (CHW)	SA-CA-GRNN (CHW)	CA-GRNN (HCBHW)	SA-CA-GRNN (HCBHW)
R^2	0.999	0.999	0.968	0.975
RMSE (kg pc)	19.93	19.93	0.31	0.27
MAE (kg pc)	9.16	9.07	0.20	0.18
MAPE (%)	9.73	9.05	12.60	11.80
IA	0.998	0.998	0.990	0.992
FA1.1 (%)	56.52	65.22	47.37	47.37
FA1.2 (%)	91.30	91.30	78.95	84.21
FA1.25 (%)	100	100	89.47	100

Although both models for the prediction of CHW generation (Table 5) demonstrated good performance, better predictions were achieved with the SA-CA-GRNN (CHW) model, which had MAPE = 9.05% and FA1.1 = 65.22%.

In the case of HCBHW modelling, similar conclusions were observed: SA-CA-GRNN (HCBHW) model has slightly better results for all prediction parameters than CA-GRNN (HCBHW) (Table 5).

The results obtained for the test data using the CA-GRNN (CHW) and SA-CA-GRNN (CHW) models for the prediction of CHW generation and for the test data using CA-GRNN (HCBHW) and SA-CA-GRNN (HCBHW) models are presented in Fig. 2.

It should also be noted that the high value of the coefficient of determination for the prediction of CHW has also been influenced by extremely high value of CHW generation in Estonia. The annual amount of generated CHW in Estonia is about 25 times greater than the average annual amount of generated CHW in the other analysed countries. If the data for Estonia were not been used, the value of the coefficient of determination would have been 0.955, which can still be considered as high. In that case, MAPE, IA, FA1.1 and FA1.2 would be just slightly changed (MAPE = 9.16%, IA = 0.983; FA1.1 = 63.64% and FA1.2 = 90.91%), which would not significantly affect the quality of the results obtained from the model.

The most prominent difference between the actual and predicted value for CHW modelling is obtained for Norway. This can be explained by the rapid increase of the amount of CHW generation in 2012 and 2014 [8].

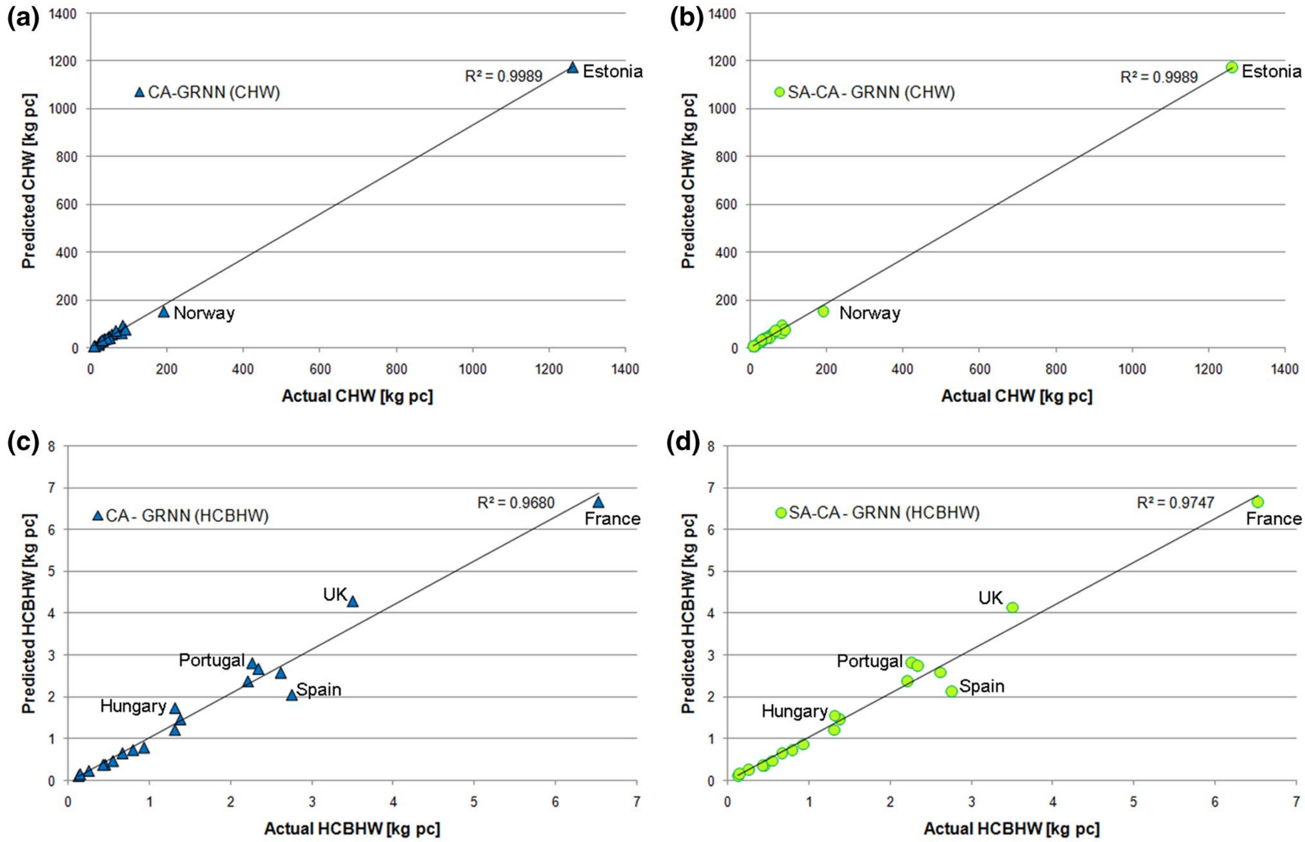


Fig. 2 Comparison of the actual data and model results for annual CHW (**a** and **b**, respectively) and HCBHW generation (**c** and **d**)

The highest rate of HCBHW generation is observed in France, but the difference between France and the other analysed countries is less prominent than in the case of generation of CHW. The deviation between the actual data and predicted values for some countries (Hungary and Portugal) in HCBHW modelling results can be related to the uncertainty of the data used, since the values for these countries were estimated in 2004, 2006 and 2008 [8]. On the other hand, the deviation between the actual and predicted HCBHW values for UK and Spain are probably a consequence of the variations in healthcare waste management legislation and policies across the UK [54] and between the Spanish Autonomous Communities [55].

Further analysis of the performance of the models can be made by assessing the discrepancy ratio (Fig. 3).

As can be seen from Fig. 3 and Table 5, the prediction of the annual amount of CHW with the relative error within ± 10% is increased from 56.6% for CA-GRNN (CHW) to 65.2% for SA-CA-GRNN (CHW) model. There is no prediction with the relative error greater than 25% for either of the models.

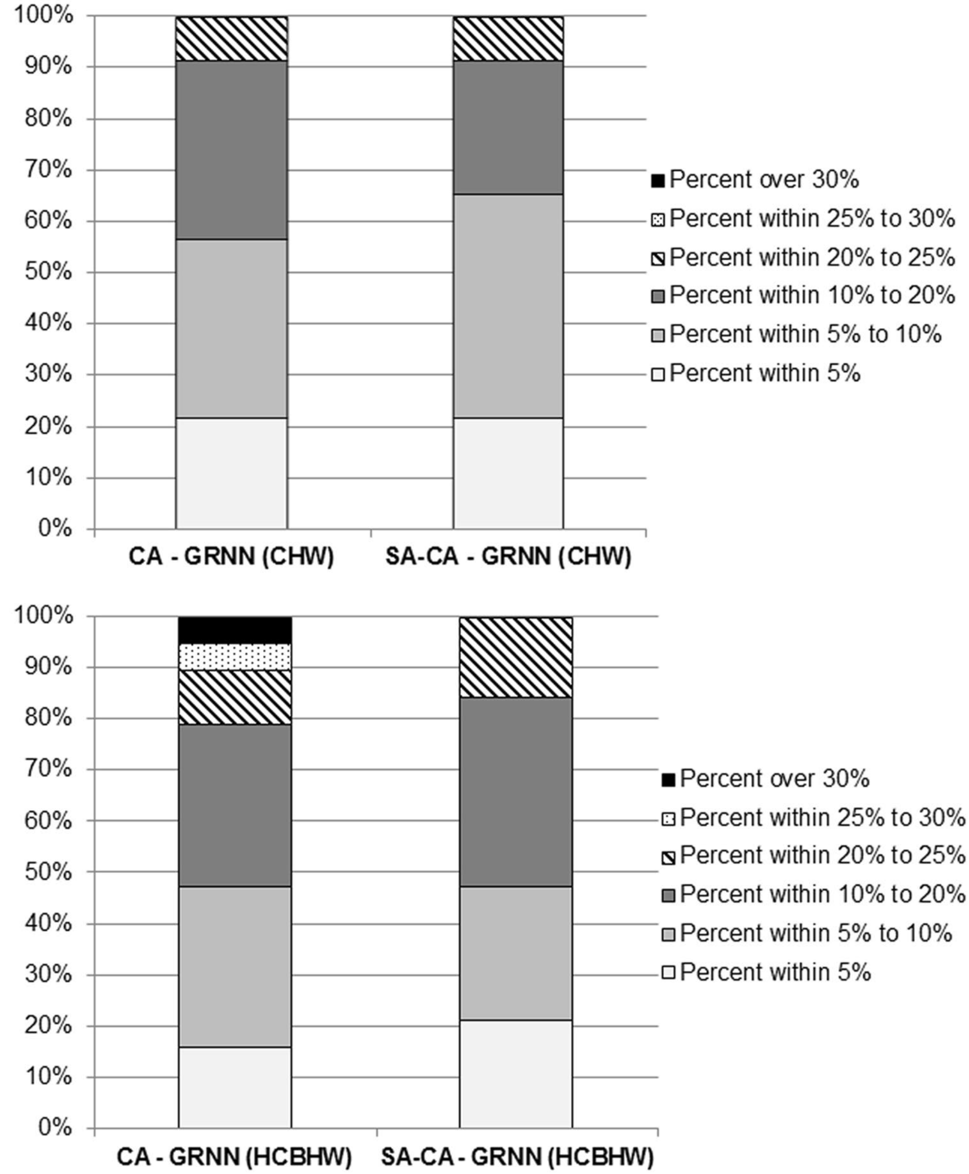
The main advantage of the SA-CA-GRNN (CHW) model lies in its improved predicting capabilities; even though the model was trained with the number of inputs reduced by 45% compared to the initial dataset (i.e. 5 inputs from initial 11 are removed).

Although the predictions of annual amount of HCBHW with the relative error within ± 10% are the same for both models (47.37%), the main advantage of the SA-CA-GRNN (HCBHW) model over the CA-GRNN (HCBHW) model is that there are no predictions with errors greater than 25%, while for 10.5% predictions of CA-GRNN (HCBHW) the error is greater than 25% (Fig. 3). Also, it should be noted that the SA-CA-GRNN (HCBHW) model, despite its better prediction capabilities, was trained with 23% less input data than the CA-GRNN (HCBHW) model.

If the performance of CHW and HCBHW models is compared with the literature available for solid waste and healthcare waste models (Table 6), it can be observed that this data-driven methodology has been successfully extended to the prediction of CHW and HCBHW at the national level, which was the main aim of this study.

The prediction obtained by calculation of annual generation of the hazardous hospital waste (infectious and biological) in Iran for 2008 [26] (Table 6) amounted to 482,965 t or 6.63 kg per capita, which is significantly higher than the HCBHW generation rates in the European countries (Table 2).

Fig. 3 Discrepancy ratios: CA-GRNN (CHW) and SA-CA-GRNN (CHW) models (figure above) and CA-GRNN (HCBHW) and SA-CA (HCBHW) models (figure below)



Considering the short period of collecting data, starting from 2004, it is assumed that in coming years, when these datasets become more uniform and more reliable, the accuracy of the predicting models will increase and opportunities for their application will be better.

Conclusion

The aim of this research was to develop a model for the prediction of annual amount of generated hazardous waste in European countries using general regression neural network (GRNN). From highly heterogeneous hazardous waste, chemical hazardous waste (CHW) and healthcare and biological hazardous waste (HCBHW), which jointly make up about 40% of total hazardous waste, were selected as the output parameters. Two separate GRNN models were created for the prediction of annual amounts of generated CHW and HCBHW.

Different social, economic, industrial and agricultural indicators were used as the input values, which were selected using correlation analysis (CA) and sensitivity analysis (SA). The GRNN architectures were trained with the biannual data from EU countries for the period 2004–2012 and tested with the data from 2014.

Performance of the developed prediction models were evaluated using: the coefficient of determination (R^2), the root mean squared error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE) and the percentage of prediction within a factor of 1.1 and 1.2 of observed values (FA1.1 and FA1.2).

Table 6 Methodology and the performance of models reported for solid and healthcare waste prediction

Authors	Level	Methodology	Performance (R^2 value)
Solid waste prediction			
Antanasijević et al. [19]	National	Standard back-propagation ANN and GRNN were used to predict municipal waste generation (MWG) in EU countries based on basic sustainability parameters	(GRNN) R^2 from 0.80 to 0.84[a]
Adamović et al. [20]	National	GRNN optimized using structural brake analysis was applied for the prediction of MSW generation. As inputs were used different indicators related to economy, demography, industry and environment, as well as those that describe social and consumer habits	(GRNN) R^2 = 0.96; MAPE = 4.0%; FA (1.1) = 86.4%
Noori et al. [56]	Local	ANN models optimized by principal component analysis and gamma test (GT) technique were utilized for weekly solid waste prediction in Mashhad (Iran)	(GT-ANN) R^2 = 0.81
Rimaitytė et al. [57]	Local	The combination of autoregressive integrated moving average (ARIMA) and seasonal exponential smoothing (SES) techniques were found to be the most accurate for MSW forecasting in Kaunas (Lithuania)	(ARIMA + SES) R^2 = 0.55; MAPE = 8.92%
Healthcare waste			
Sabour et al. [26]	National	Linear regression model has been proposed for the generation of hospital waste in Iran based on the number of hospitals and active beds	Validation was not reported
Jahandideh et al. [27]	Regional	ANN and multiple linear regression (MLR) models were used to predict medical waste generation in Fars province (Iran) based on hospital related parameters	(ANN) R^2 = 0.99
Eleyan et al. [28]	Regional	System dynamics model was developed for the prediction of medical solid waste generation rates in Jenin District (Palestine) based on the characteristics of hospitals	Validation was not reported
Al-Khatib et al. [29]	Local	MLR model for hospital solid waste generation are proposed based on the characteristics of hospitals in Nablus city (Palestine)	Hazardous hospital solid waste R^2 = 0.984
Karpušenkaitė et al. [30]	Regional and national	ANN, MLR, support vector machines and different non-parametric regression methods were applied for the forecasting of medical waste generation in Lithuania using diverse set of inputs	Regional level (GANPR)[b] R^2 = 0.905; MAPE = 31.90% National level (SSNPR)[c] R^2 = 0.986; MAPE = 8.59%

[a]Depending on test set

[b]Generalised additive non-parametric regression

[c]Smoothing splines non-parametric regression

The model for prediction of annual quantities of generated CHW expressed in kg per capita covered 23 European countries. Using correlation analysis, and then sensitivity analysis, the number of inputs was reduced from initial 11 to 6, and the performance of the model was improved.

For the model of the annual quantities of generated HCBHW (also in kg per capita) available datasets for 19 European countries were used. The best modelling results were achieved using sensitivity analysis, whereby the number of input variables was reduced from 16 to 10. The main advantage of the optimized model is that 84.2% of predicted values had an error less than 20% and that there were no predictions with the error greater than 25%.

Beside for the prediction of annual generation of CHW and HCBHW, the developed GRNN models can also be applied for the simulation of various scenarios that include possible regulatory, industrial and agricultural changes and measures, and in that way support the decision-making process on sustainable development at the regional or national level.

Acknowledgements The authors are grateful to the Ministry of Education, Science and Technological Development of the Republic of Serbia, Project No. 172007 for financial support.

Appendix A. Details on the selected Inputs

Gross domestic product (GDP) at market prices expressed in purchasing power standard (PPS) eliminates the differences in price levels among countries, and allows the comparison of economies significantly different in absolute size. PPS is an artificial currency unit and theoretically one PPS can buy the same amount of goods and services in each country [58, 59].

Domestic material consumption (DMC) represents the sum of the annual amount of raw material extracted from the country and the difference of all physical imports and exports [60].

Human development index (HDI) is a composite index that measures the average achievement in each specific country in three fundamental dimensions of human development: long and healthy life, knowledge and decent standard of living. The indicators used to calculate HDI are: for the health component—life expectancy at birth, for the education component—mean years of schooling (for adults) and expected years of mean schooling (for children), and for the standard of living component—the log of the purchasing power parity (PPP)—adjusted Gross National Income (GNI) [61, 62].

Value added of industry (or agriculture) refers to the contribution of industry (or agriculture) to overall GDP. Value added is the net output of a sector (e.g. industry or agriculture) adding up all outputs and subtracting intermediate inputs. Value added in industry comprises value added in mining, manufacturing, construction, electricity, water and gas. Value added in agriculture comprises value added in livestock production and cultivation of crops, but also includes value added in forestry, hunting and fishing [63].

Final energy consumption includes information about annual data on quantities of crude oil, oil products, natural gas and manufactured gases, electricity and derived heat, solid fossil fuels, renewables and wastes covering the full spectrum of the energy sector from supply through transformation to final energy consumption by chemical and petrochemical industries [64, 65]. Final energy consumption in chemical and petrochemical industries for this study is normalized and expressed as kg of oil equivalent per capita.

Total intramural expenditure on research and development is in fact gross domestic expenditure on research and development in the specific country. It aggregates the total expenditure on research and development (RD) financed by a country's institutions including research and development performed abroad but financed by national institutions—while it excludes RD performed within a country, but funded from abroad [66]. The funding sectors are: the business enterprise sector, the government sector, public general university funds, higher education sector and private non-profit sector.

Life expectancy at birth is an indicator that represents the average number of years that a new-born could expect to live if the prevailing patterns of mortality at the time of its birth were to stay the same throughout its life [67].

Data about the inability to face unexpected financial expenses were obtained through householder surveys on the capacity to cover unexpected cost from own resources. Unexpected costs refer to amounts equal to the poverty threshold for each specific country expressed as a monthly sum [68, 69].

Death rate, crude or crude death rate indicates the number of deaths occurring during the year, per 1000 inhabitants estimated at midyear [63].

Bed days indicate the number of hospitals days per capita for all causes of diseases except V00-Y98 (External causes of morbidity and mortality) and Z38 (live born infants according to place of birth) [70].

Appendix B. The model performance metrics used in this study

The coefficient of determination represents the degree of linear correlation between the predicted and observed values. This is a measure of the ability of the model to predict an outcome in the linear regression setting [71]. Although it provides limited information, R^2 is often used due to its simplicity.

If O_i is the observed value, P_i is the predicted value, O_m and P_m are the means of the observed and predicted data, and n is the number of observations then the coefficient of determination can be shown as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - O_m)^2}. \quad (7)$$

If there is an intercept in the linear model, R^2 is equal to the square of the correlation coefficient [71]:

$$R^2 = \left(\frac{\sum_{i=1}^{n} (O_i - O_m)(P_i - P_m)}{\sqrt{\sum_{i=1}^{n} (O_i - O_m)^2 \sum_{i=1}^{n} (P_i - P_m)^2}} \right)^2. \quad (8)$$

The root mean squared error (RMSE) is the square root of the mean square error (MSE) which is the average of the squares of the difference between the actual and predicted values [72]. Due to the fact that the RMSE is expressed in the units of the variables, it is much more convenient and easier to interpret than the coefficient of determination [47],

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\left[\sum_{i=1}^{n}(P_i - O_i)^2\right]}. \tag{9}$$

The mean absolute error (MAE) is the mean value of the absolute value of the difference between the observed and predicted values:

$$\mathrm{MAE} = \sum_{i=1}^{n}\frac{|O_i - P_i|}{n}. \tag{10}$$

The mean absolute percentage error (MAPE) represents a measure of the accuracy of the predicted values in relation to the measured, expressed in percentages:

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|O_i - P_i|}{|O_i|}. \tag{11}$$

Generally, when MAPE is less than 10%, the forecast is highly accurate, at 10–20% the forecast is good, at 20–50% the forecast is reasonable, and for more than 50%, the forecast is inaccurate [73], but the limits used for the assessment of forecast accuracy may vary for different applications.

The percentage of predictions within a factor of the observed values (FA) demonstrates the ability of the model to give accurate prediction for each test case [47],

$$\mathrm{FA1.1} = 0.9 < \frac{P_i}{O_i} < 1.1, \tag{12}$$

$$\mathrm{FA1.2} = 0.83 < \frac{P_i}{O_i} < 1.2, \tag{13}$$

$$\mathrm{FA1.25} = 0.8 < \frac{P_i}{O_i} < 1.25. \tag{14}$$

The index of agreement shows the degree of medium agreement between the actual and the predicted values, as well as the error between them. It is a relative and standardized statistical performance indicator with values between 0 (poor agreement) and 1 (excellent agreement) [74],

$$\mathrm{IA} = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}\left(|P_i - O_m| + |O_i - O_m|\right)^2}. \tag{15}$$

References

1. European Parliament and the Council of the European Union (2008) Directive 2008/98/EC of the European Parliament and of the Council on waste and repealing certain directives (Waste framework). OJ L 312,22.11.2008, 3–30. doi:2008/98/EC.; 32008L0098
2. Kang Y-Y, Jeon T-W, Kim W-I, Shin SK, Yeon J-M, Somasundaram S (2014) Correlation study of hazardous waste characteristics among various chemical processes in Republic of Korea. J Mater Cycles Waste Manag 16(3):566–575. https://doi.org/10.1007/s10163-013-0208-8
3. Court CD, Munday M, Roberts A, Turner K (2015) Can hazardous waste supply chain "hotspots" be identified using an input-output framework? Eur J of Oper Res 241(1):177–187. https://doi.org/10.1016/j.ejor.2014.08.011
4. Schuhmacher M, Fàbrega F, Kumar V, García F, Nadal M, Domingo JL (2014) A PBPK model to estimate PCDD/F levels in adipose tissue: Comparison with experimental values of residents near a hazardous waste incinerator. Environ Int 73:150–157. https://doi.org/10.1016/j.envint.2014.07.020
5. Zakaria B, Abdullah R, Ramli MF, Latif PA (2012) Selection criteria using the Delphi method for siting an integrated hazardous waste disposal facility in Malaysia. J Environ Plann Manag 56(4):1–19. https://doi.org/10.1080/09640568.2012.689614
6. United Nations (2007) Indicators of Sustainable Development: Guidelines and Methodologies. Third Edit. United Nations publication, New York, USA. http://www.un.org/esa/sustdev/natlinfo/indicators/guidelines.pdf
7. van Beusekom C (1999) Hazardous waste in the European Union. Statistics in focus—Environment and energy, Luxembourg: Eurostat 1–7. http://edz.bib.uni-mannheim.de/www-edz/pdf/statinf/99/CA-NQ-99-007-EN-I-EN.pdf. Accessed 22 Mar 2016
8. European Commission Eurostat (2017) Waste statistics—statistics explained. http://ec.europa.eu/eurostat/statistics-explained/index.php/Waste_statistics. Accessed 22 Mar 2016
9. EPA US (2012). Chemical safety for sustainability chemical, strategic research action plan 2012–2016. EPA 601/R-12/006:pp 1–64. http://www.epa.gov/sites/production/files/2014-06/documents/css-strap.pdf. Accessed 4 May 2016
10. Rudel RA, Attfield KR, Schifano JN, Brody JG (2007) Chemicals causing mammary gland tumors in animals signal new directions for epidemiology, chemicals testing, and risk assessment for breast cancer prevention. Cancer 109(12):2635–2666. https://doi.org/10.1002/cncr.22653
11. Komilis D, Fouki A, Papadopoulos D (2012) Hazardous medical waste generation rates of different categories of healthcare facilities. Waste Manag 32(7):1434–1441. https://doi.org/10.1016/j.wasman.2012.02.015
12. Birpinar ME, Bilgili MS, Erdoğan T (2009) Medical waste management in Turkey: a case study of Istanbul. Waste Manag 29(1):445–448. https://doi.org/10.1016/j.wasman.2008.03.015
13. Mmereki D, Baldwin A, Li B, Liu M (2017) Healthcare waste management in Botswana: storage, collection, treatment and disposal system. J Mater Cycles Waste Manag 19(1):351–365. https://doi.org/10.1007/s10163-015-0429-0
14. Rushbrook P, Zghondi R (2005) Better healthcare waste management: an integral component of health investment. World Health Organization, Amman http://apps.who.int/iris/bitstream/10665/119762/1/dsa515.pdf. Accessed 22. Mar 2016
15. Shannon AL, Woolridge A (2011) Chap. 23 - Medical Waste. Elsevier Inc. https://doi.org/10.1016/B978-0-12-381475-3.10023-3
16. Sartaj M, Arabgol R (2015) Assessment of healthcare waste management practices and associated problems in Isfahan Province (Iran). J Mater Cycles Waste Manag 17(1):99–106. https://doi.org/10.1007/s10163-014-0230-5
17. Elimelech E, Ayalon O, Flicstein B (2011) Hazardous waste management and weight-based indicators - The case of Haifa

Metropolis. J Hazard Mater 185(2–3):626–633. https://doi.org/10.1016/j.jhazmat.2010.09.064
18. Specht DF (1991) A general regression neural network. IEEE Trans Neural Netw 2(6):568–576. https://doi.org/10.1109/72.97934
19. Antanasijević D, Pocajt V, Popović I, Redžić N, Ristić M (2013) The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. Sustain Sci 8(1):37–46. https://doi.org/10.1007/s11625-012-0161-9
20. Adamović V, Antanasijević D, Ristić M, Perić-Grujić A, Pocajt V (2017) Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis. Environ Sci Pollut Res 24(1):299–311. https://doi.org/10.1007/s11356-016-7767-x
21. Statistics Norway (2016) Hazardous waste, 2015. Oslo. http://www.ssb.no/en/natur-og-miljo/statistikker/spesavf/aar. Accessed 18 May 2017
22. Gusca J, Kalnins SN, Blumberga D, Bozhko L, Khabdullina Z, Khabdullin A (2015) Assessment method of health care waste generation in Latvia and Kazakhstan. Energy Proc 72:175–179. https://doi.org/10.1016/j.egypro.2015.06.025
23. Ben-Gal I (2005) Outlier Detection. In: Maimon O, Rockach L (eds). Data mining and knowledge discovery handbook: a complete guide for practitioners and researchers. Kluwer Academic Publishers pp 131–146. https://doi.org/10.1007/0-387-25465-x_7
24. Liu H, Shah S, Jiang W (2004) On-line outlier detection and data cleaning. Comput Chem Eng 28(9):1635–1647. https://doi.org/10.1016/j.compchemeng.2004.01.009
25. Beigl P, Lebersorger S, Salhofer S (2008) Modelling municipal solid waste generation: a review. Waste Manag 28(1):200–214. https://doi.org/10.1016/j.wasman.2006.12.011
26. Sabour MR, Mohamedifard A, Kamalan H (2007) A mathematical model to predict the composition and generation of hospital wastes in Iran. Waste Manag 27(4):584–587. https://doi.org/10.1016/j.wasman.2006.05.010
27. Jahandideh S, Jahandideh S, Asadabadi EB, Askarian M, Movahedi MM, Hosseini S et al (2009) The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation. Waste Manag 29(11):2874–2879. https://doi.org/10.1016/j.wasman.2009.06.027
28. Eleyan D, Al-Khatib I, Garfield J (2013) System dynamics model for hospital waste characterization and generation in developing countries. Waste Manag Res 31(10):986–995. https://doi.org/10.1177/0734242X13490981
29. Al-Khatib IA, Abu Fkhidah I, Khatib JI, Kontogianni S (2016) Implementation of a multi-variable regression analysis in the assessment of the generation rate and composition of hospital solid waste for the design of a sustainable management system in developing countries. Waste Manag Res 34(3):225–234. https://doi.org/10.1177/0734242X15622813
30. Karpušenkaitė A, Ruzgas T, Denafas G (2016) Forecasting medical waste generation using short and extra short datasets: case study of Lithuania. Waste Manag Res 34(4):378–387. https://doi.org/10.1177/0734242X16628977
31. Granados AJ, Peterson PJ (1999) Hazardous waste indicators for national decision makers. J Environ Manage 55(4):249–263. https://doi.org/10.1006/jema.1999.0254
32. Tripathy M (2010) Power transformer differential protection using neural network Principal Component Analysis and Radial Basis Function Neural Network. Simul Model Pract Theory 18(5):600–611. https://doi.org/10.1016/j.simpat.2010.01.003
33. Mustapha A, Aris AZ, Juahir H, Ramli MF, Kura NU (2013) River water quality assessment using environmentric techniques: case study of Jakara River Basin. Environ Sci Pollut Res 20(8):5630–5644. https://doi.org/10.1007/s11356-013-1542-z
34. Hamilton LC (1991) Modern data analysis: a first course in applied statistics. Technometrics 33(4):487–488
35. Walczak S, Cerpa N (1999) Heuristic principles for the design of artificial neural networks. Inf Softw Technol 41(2):107–117. https://doi.org/10.1016/S0950-5849(98)00116-5
36. Gheyas IA, Smith LS (2010) A neural network-based framework for the reconstruction of incomplete data sets. Neurocomputing 73(16–18):3039–3065. https://doi.org/10.1016/j.neucom.2010.06.021
37. Tomandl D, Schober A (2001) A modified general regression neural network (MGRNN) with new, efficient training algorithms as a robust "black box"-tool for data analysis. Neural Netw 14(8):1023–1034. https://doi.org/10.1016/S0893-6080(01)00051-X
38. Zhou Q, Jiang H, Wang J, Zhou J (2014) A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. Sci Total Environ 496(1):264–274. https://doi.org/10.1016/j.scitotenv.2014.07.051
39. Gheyas IA, Smith LS (2011) A novel neural network ensemble architecture for time series forecasting. Neurocomputing 74(18):3855–3864. https://doi.org/10.1016/j.neucom.2011.08.005
40. Millie DF, Weckman GR, Young W, Ivey JE, Carrick HJ, Fahnenstiel GL (2012) Modeling microalgal abundance with artificial neural networks: demonstration of a heuristic "Grey-Box" to deconvolve and quantify environmental influences. Environ Model Softw 38(1):27–39. https://doi.org/10.1016/j.envsoft.2012.04.009
41. Kisi O (2006) Generalized regression neural networks for evapotranspiration modelling. J Hydrol Sci 51(6):1092–1105. https://doi.org/10.1623/hysj.51.6.1092
42. Palani S, Liong SY, Tkalich P (2008) An ANN application for water quality forecasting. Mar Poll Bull 56(9):1586–1597. https://doi.org/10.1016/j.marpolbul.2008.05.021
43. Šiljić A, Antanasijević D, Perić-Grujić A, Ristić M, Pocajt V (2014) Artificial neural network modelling of biological oxygen demand in rivers at the national level with input selection based on Monte Carlo simulations. Environ Sci Pollut Res 22(6):4230–4241. https://doi.org/10.1007/s11356-014-3669-y
44. Yaseen ZM, El-shafie A, Jaafar O, Afan HA, Sayl KN (2015) Artificial intelligence based models for stream-flow forecasting: 2000–2015. J Hydrol 530:829–844. https://doi.org/10.1016/j.jhydrol.2015.10.038
45. Specht DF (1990) Probabilistic neural networks. Neural Netw 3(1):109–118. https://doi.org/10.1016/0893-6080(90)90049-Q
46. Sawant SS, Topannavar PS (2015) Introduction to probabilistic neural network—used for image classifications. Int J Adv Res Comput Sci Softw Eng 5(4):279–283
47. Antanasijević D, Ristić M, Perić-Grujić A, Pocajt V (2014) Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis. Int J Greenh Gas Control 20:244–253. https://doi.org/10.1016/j.ijggc.2013.11.011
48. Sözen A, Gülseven Z, Arcaklioğlu E (2007) Forecasting based on sectoral energy consumption of GHGs in Turkey and mitigation policies. Energy Policy 35(12):6491–6505. https://doi.org/10.1016/j.enpol.2007.08.024
49. Pahlavan R, Omid M, Akram A (2012) Energy input-output analysis and application of artificial neural networks for predicting greenhouse basil production. Energy 37(1):171–176. https://doi.org/10.1016/j.energy.2011.11.055
50. Kialashaki A (2014) Evaluation and Forecast of Energy Consumption in Different Sectors of the United States Using Artificial Neural Networks. Theses and Dissertations. Paper 628. University of Wisconsin-Milwaukee
51. Jacobs G, Šlaus I (2010) Indicators of economic progress: the power of measurement and human welfare. CADMUS J 1(1):53–113

52. Costanza R, Hart M, Posner S, Talberth J (2009) Beyond GDP: the need for new measures of progress.The Pardee papers no. 4 - January 2009. Boston University, Boston, MA, USA. https://www.bu.edu/pardee/files/documents/PP-004-GDP.pdf
53. Dogan E, Sengorur B, Koklu R (2009) Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. J Environ Manage 90(2):1229–1235. https://doi.org/10.1016/j.jenvman.2008.06.004
54. CIWM (2014) An introductory guide to healthcare waste management in England and Wales. Northampton, UK. https://www.ciwm-journal.co.uk/downloads/Healthcare-Waste-WEB.pdf. Accessed 8 Apr 2018
55. Insa E, Zamorano M, López R (2010) Critical review of medical waste legislation in Spain. Resour Conserv Recycl 54(12):1048–1059. https://doi.org/10.1016/j.resconrec.2010.06.005
56. Noori R, Karbassi A, Salman Sabahi M (2010) Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. J Environ Manage 91:767–771. https://doi.org/10.1016/j.jenvman.2009.10.007
57. Rimaityte I, Ruzgas T, Denafas G, Racys V, Martuzevicius D (2012) Application and evaluation of forecasting methods for municipal solid waste generation in an eastern-European city. Waste Manag Res 30:89–98. https://doi.org/10.1177/0734242X10396754
58. Eurostat (2016) Glossary: purchasing power standard (PPS)—statistics explained. http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Purchasing_power_standard. Accessed 20 Apr 2016
59. NSI (2016) GDP at Market prices per inhabitant in Euro and in PPS | National statistical institute Republic of Bulgaria. http://www.nsi.bg/en/content/5226/gdp-market-prices-inhabitant-euro-and-pps. Accessed 20 Apr 2016
60. Weisz H, Krausmann F, Amann C, Eisenmenger N, Erb KH, Hubacek K, Fisher-Kowalski (2006) The physical economy of the European Union: cross-country comparison and determinants of material consumption. Ecol Econ 58(4):676–698. https://doi.org/10.1016/j.ecolecon.2005.08.016
61. Harttgen K, Klasen S (2012) A household-based human development index. World Dev 40(5):878–899. https://doi.org/10.1016/j.worlddev.2011.09.011
62. UNDP (2015) Human development report 2015. Work for human development. New York http://hdr.undp.org/sites/default/files/2015_human_development_report.pdf. Accessed 5 May 2016
63. The World Bank (2014) World Development Indicators 2014. Washington: International Bank for Reconstruction and Development/The World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/18237/9781464801631.pdf?sequence=1&isAllowed=y. Accessed 5 May 2016
64. Saygin D, Patel M, Tam C, Gielen D (2009) Chemical and Petrochemical sector. Potential of best practice technology and other measures for improving energy efficiency. IEA Information Paper OECD/IEA, pp 1–60
65. Eurostat (2016) Simplified energy balances - annual data http://ec.europa.eu/eurostat/web/products-datasets/-/nrg_100a. Accessed 5 May 2016
66. OECD (2002) Proposed standard practice for surveys of research and experimental development: "Frascati Manual 2002", the measurement of scientific and technological activities. OECD Publications, Paris
67. The World Bank (2016) Life expectancy at birth. http://data.worldbank.org/indicator/SP.DYN.LE00.IN. Accessed 6 May 2016
68. Ward T (2008) Material deprivation in the EU (Chapter 5). In: Ward T, Lelkes O, Sutherland H, Tóth I G (eds) Social inclusion and income distribution in the European Union - 2008, European Commission, Tárki, Budapest. pp 114–128. http://ec.europa.eu/social/BlobServlet?docId=3991&langId=en
69. Eurostat (2016) Inability to face unexpected financial expenses (source: SILC) - Eurostat. http://ec.europa.eu/eurostat/en/web/products-datasets/-/ILC_MDES04. Accessed 6 May 2016
70. Eurostat (2016) SCL—International statistical classification of diseases and related health problems (ICD-10 2007). RAMON-Reference and Management of Nomenclatures. http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL_LINEAR&IntCurrentPage=1&StrNom=CL_ICD10&StrLanguageCode=EN. Accessed 17 Oct 2016
71. Renaud O, Victoria-Feser MP (2010) A robust coefficient of determination for regression. J Stat Plan Inference 140(7):1852–1862. https://doi.org/10.1016/j.jspi.2010.01.008
72. Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. Geosci Model Dev 7(3):1247–1250. https://doi.org/10.5194/gmd-7-1247-2014
73. Pao H-T, Fu H-C, Tseng C-L (2012) Forecasting of CO_2 emissions, energy consumption and economic growth in China using an improved grey model. Energy 40(1):400–409. https://doi.org/10.1016/j.energy.2012.01.037
74. Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. Int J Climatol 32(13):2088–2094. https://doi.org/10.1002/joc.2419