Environ Sci Pollut Res (2015) 22:18849–18858
DOI 10.1007/s11356-015-5075-5

RESEARCH ARTICLE

Modeling of ammonia emission in the USA and EU countries using an artificial neural network approach

Lidija J. Stamenković[1] · **Davor Z. Antanasijević**[2] · **Mirjana Đ. Ristić**[1] · **Aleksandra A. Perić-Grujić**[1] · **Viktor V. Pocajt**[1]

Received: 19 February 2015 / Accepted: 14 July 2015 / Published online: 24 July 2015

Abstract Ammonia emissions at the national level are frequently estimated by applying the emission inventory approach, which includes the use of emission factors, which are difficult and expensive to determine. Emission factors are therefore the subject of estimation, and as such they contribute to inherent uncertainties in the estimation of ammonia emissions. This paper presents an alternative approach for the prediction of ammonia emissions at the national level based on artificial neural networks and broadly available sustainability and economical/agricultural indicators as model inputs. The Multilayer Perceptron (MLP) architecture was optimized using a trial-and-error procedure, including the number of hidden neurons, activation function, and a back-propagation algorithm. Principal component analysis (PCA) was applied to reduce mutual correlation between the inputs. The obtained results demonstrate that the MLP model created using the PCA transformed inputs (PCA-MLP) provides a more accurate prediction than the MLP model based on the original inputs. In the validation stage, the MLP and PCA-MLP models were tested for ammonia emission predictions for up to 2 years and compared with a principal component regression model. Among the three models, the PCA-MLP demonstrated the best performance, providing predictions for the USA and the majority of EU countries with a relative error of less than 20 %.

Responsible editor: Gerhard Lammel

✉ Davor Z. Antanasijević
dantanasijevic@tmf.bg.ac.rs

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

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

Keywords ANN · MLP · PCA · National emissions · Ammonia emissions

Introduction

Gaseous ammonia (NH_3) is the most abundant alkaline gas in the atmosphere (Behera et al. 2013) with a wide variety of impacts, including photochemical air pollution, reduced visibility, changes in biodiversity, and stratospheric ozone depletion (Beusen et al. 2008). Ammonia is emitted into the atmosphere from various sources: agriculture, road transport, industrial processes, etc., and recent studies suggest that human activities accelerate the production of reactive nitrogen on a global scale (Aneja et al. 2008). Agricultural activities are by far the largest global source of NH_3 to the atmosphere (Gruber and Galloway 2008). Estimates show that 83 % of the nitrogen used as fertilizer is lost to the environment (Erisman et al. 2008), while in the USA, emissions of NH_3 account for 30 % of these losses (Houlton et al. 2013). The agriculture sector is responsible for over 90 % of NH_3 emissions in Europe (European Environment Agency (EEA) 2014). Emission estimates from Southeast Asia (e.g., China and India) and other developing countries are very limited (Aneja et al. 2008).

Generally, the main source of air pollutant emission data, including NH_3, is emission inventories, e.g., USEPA National Emission Inventory (NEI) for the USA and the EMEP/EEA Air Pollutant Emission Inventory Guidebook for the EU countries. The value of emissions estimated by models based on emission inventories are a compilation of a large number of input parameters (over 300). These parameters include, among

others, uncontrolled emission factors, removal efficiency of emission control measures, and the extent to which such measures are applied (Amann et al. 2011). Developing accurate emission inventories is a difficult task because many factors are involved in the generation and dispersion of ammonia (Aneja et al. 2008). Detailed models for the estimation of national NH_3 emission (Gilliland et al. 2006; Paulot et al. 2014) and the estimation of NH_3 emission from agricultural systems, e.g., livestock operations (Pinder et al. 2004; Webb and Misselbrook 2004; Li et al. 2012; Barrancos et al. 2013; Rotz et al. 2014) and mineral fertilizers (Cooter et al. 2012), have been developed. Many of those models require a detailed knowledge of local environmental conditions and farming practices which in the most part, is generally not available (Paulot et al. 2014).

One of possible alternatives for modeling can be Artificial Neural Networks (ANNs), which are adaptable systems that can determine relationships between different sets of data (Karacan 2008). ANNs can be visualized as a black-box, non-linear and non-parametric regression method, with good flexibility, since they do not need a rigid mathematical model and because the calibration parameters can be determined using data, through a training or learning step (Despagne and Massart 1998). After training, the neural network is able to generalize its knowledge and make predictions for an unknown example (Despagne and Massart 1998; Raimundo and Narayanaswamy 2001). ANNs were successfully applied for the prediction of NH_3 concentration in the air (Raimundo and Narayanaswamy 2001) and NH_3 emissions from a field-applied manure (Lim et al. 2007), as well as for the estimation of the annual emissions of other air pollutants, e.g., PM10 (Antanasijević et al. 2013) and greenhouse gases (Antanasijević et al. 2014).

This study proposes an ANN model for the prediction of NH_3 emissions at the national level, using sustainability, economical, and agricultural indicators as inputs. The main difference between inventory based models and the proposed ANN approach is that the ANN model requires substantially smaller number of inputs, and that all of them are widely available for developed and developing countries.

In this paper, the Multilayer Perceptron (MLP) model was developed using data for the USA and 21 European countries. The architectural and training parameters of the MLP model were optimized during its development. Additionally, principal component analysis (PCA) was applied for input transformation as the method for reduction of correlation between inputs. The results of the MLP models are compared with a conventional principal component regression (PCR) model, using multiple performance criteria.

Materials and methods

Input and output data

Several studies have already shown that the emission of gaseous pollutants at the national level can be successfully predicted using an approach which combines ANNs and widely available sustainability, industrial, and economic indicators (Radojević et al. 2013; Antanasijević et al. 2014). Following this approach, four indicators related to ammonia emission have been selected as inputs (Table 1.). Gross domestic product (GDP) per capita is a broadly available economy size indicator, which is selected as it indicates the general level of development for each particular country. Since NH_3 is dominantly emitted from the agriculture sector, two indicators (the number of cattle and consumption of nitrogen fertilizers) related to the production of animal wastes and to the use of nitrogen fertilizers, which are the two major sources of ammonia emission of all agricultural activities, were also selected. The transportation sector is also a considerable source of NH_3 emissions, which occurs mainly due to the use of three-way catalysts in petrol cars (Sundvor et al. 2013). Ammonia emissions that came from the transportation sector are "covered" using the indicator final energy consumption by transport (FECT), since it summarizes the use of energy products in all types of transportation, i.e., rail, road, international and domestic air transport, and inland navigation/coastal shipping (Eurostat 2014a).

The selected inputs were normalized per capita and/or per average GDP value of EU27, in order to enable the comparison of countries of different sizes. The input and output data for every European country included into this study has been published by Eurostat (Eurostat 2014b), while the data for the USA has been published by the National Agricultural Statistics Service (US Department of Agriculture (USDA) 2014), US Energy Information Administration (USEIA) (2014), and US Environmental Protection Agency (EPA) (2014). Descriptive statistics of the chosen inputs over the period 2000–2010 is presented in Table 1, while the Pearson correlation coefficients between variables and corresponding variance inflation factors (VIF) are presented in Table 2.

The NH_3 emission model was trained, validated, and tested using the available data for the USA and 21 European countries. The data from 2000 to 2008 was used for the development of the model, while the data from 2009 and 2010 was used for its validation. The descriptive statistics of NH_3 emission at the national level for each country and for different datasets are shown in Table 3.

Table 1 Input variables selected for ammonia emission modeling and their descriptive statistics

Input variable	Unit[a]	Mean	Standard deviation	Minimum	Maximum
Gross domestic product (GDP)	–[b]	0.99	0.46	0.143	1.76
Number of cattle (NC)	Cattle pc	0.19	0.15	0.051	1.43
Consumption nitrogen fertilizers (CNF)	t pc	0.03	0.05	0.006	0.51
Final energy consumption-transport (FECT)	Toe pc[c]	0.87	0.42	0.204	2.44

[a] After normalization

[b] Unitless because of normalization per GDP value of EU27

[c] The tons of oil equivalent (toe) per capita

Artificial neural networks

Multilayer Perceptron

Neural networks are universal approximators with an ability to generalize by learning non-linear relationships between presented input(s) and output(s) variable(s) (Hájek and Olej 2012). The selection of appropriate ANN architecture, number of neurons and training algorithms has a direct impact on the model performance, as it determines how the model inputs are transformed into the model outputs (Wu et al. 2014).

The MLP is a type of artificial neural network which is well-known for its ability to represent any smooth measurable functional relationship between the inputs and the outputs. The typical MLP is composed of three layers of neurons: the input layer, the hidden layer(s), and the output layer (Agirre-Basurko et al. 2006). The number of neurons in the input and output layers is equal to the number of input and output variables used in the creation of the MLP model.

One of critical issues in using MLPs is the selection of the number of neurons, and thus weights, in the hidden layers. Within the hidden layer, the inputs are summed and processed by a non-linear function, called a transfer function, or axon and there are various transfer functions. The process of finding a suitable set of weights is called "training." Training is one of the most important steps in the development of neural networks, since the weights and the network characteristics will be used later in testing data sets and making subsequent predictions (Karacan 2008). In the training phase, the objective is to find the combination of weights which results in the smallest error (Gardner and Dorling 1998). Different training algorithms could be applied to minimize the error function, among which the back-propagation (BP) algorithm, and

Table 2 Pearson correlation coefficients between variables and corresponding variance inflation factors (VIF)

	Pearson correlation				VIF
	FECT	NC	GDP	CNF	
FECT	1.000				2.06
NC	0.388[a]	1.000			2.59
GDP	0.642[a]	0.490[a]	1.000		2.92
CNF	0.065	0.136	0.108	1.000	1.52
NH_3 emission	0.420[a]	0.809[a]	0.465[a]	0.102	

[a] Correlations are significant at $p<0.01$

Table 3 Descriptive statistics of NH_3 data for the study period

Country/region	NH_3 emission (kg pc)			
	Mean	Standard deviation	Minimum	Maximum
USA	12.767	1.060	11.744	15.777
Bulgaria	7.149	0.439	6.699	7.614
Czech Republic	6.636	0.735	5.823	7.706
Denmark	15.496	1.353	13.475	17.413
Estonia	7.563	0.285	7.287	8.001
Greece	6.063	0.357	5.575	6.606
Spain	8.750	0.635	7.780	9.491
France	12.033	0.880	9.978	13.244
Cyprus	6.725	0.204	6.510	7.011
Latvia	7.259	0.309	6.846	7.711
Lithuania	9.432	1.012	8.473	10.735
Hungary	7.040	0.583	6.530	8.036
Netherlands	8.713	0.904	7.346	10.259
Austria	7.736	0.207	7.457	8.080
Poland	7.373	0.222	7.102	7.590
Portugal	5.071	0.513	4.502	5.964
Romania	8.713	0.735	7.516	9.396
Slovenia	8.833	0.270	8.530	9.246
Slovakia	4.756	0.225	4.502	4.998
Finland	7.151	0.165	6.887	7.403
UK	5.132	0.362	4.585	5.665
Dataset				
All data	8.390	3.313	4.502	26.706
Training data	8.611	3.534	4.630	26.706
Validation data	7.639	2.299	4.502	13.937

associated algorithms which are derived from BP, are most frequently used (Russo et al. 2013).

Back-propagation algorithm

The back-propagation (BP) algorithm is the most used algorithm for training the multilayer perceptron (Gardner and Dorling 1998). BP is a gradient descent algorithm, developed by Rumelhart et al. (1986), in which the network weights are moved along the negative of the gradient of the performance function. Details on BP algorithms and its implementation can be found elsewhere (Bishop 2006). The variations of the basic algorithm include heuristics and standard numerical optimization techniques (Bolanča et al. 2008):

i. Learning rate gradient descent - the weights are updated by the change dictated by learning rate and error gradient. The gradient descent algorithm with adaptive learning rate is based on a heuristic technique (Vogl et al. 1988) and is much faster than the standard steepest descent algorithm as it allows the learning rate to change during the training process. The $(k+1)$th iteration of this algorithm can be written as:

$$w_{k+1} = w_k - \alpha_k g_k \tag{1}$$

where w_k is a vector of current weights and biases, g_k is the current gradient and α_k is the learning rate.

ii. Conjugate gradient—the weights are updated not only by learning rate but also with so-called momentum. The basic BP algorithm adjusts the weights in the steepest descent direction. In conjugate gradient algorithms, a search is performed along conjugate directions, which generally produces faster convergence than the steepest descent directions (Ramesh et al. 2008):

$$w_{k+1} = w_k + \alpha_k p_k \tag{2}$$

where p_k is the current search direction. The general procedure for determining the new search direction is to combine the new steepest descent direction with the previous search direction:

$$p_k = -g_k + \beta_k p_{k-1} \tag{3}$$

where p_{k-1} is the previous search direction and β_k is constant (momentum).

iii. Newton methods—instead of using learning rate or conjugate gradient, the weights are calculated by Quasi-Newton method, using approximated Hessian matrix of the performance index at the current values of weights (A_k). The weight update for the Newton's method is:

$$w_{k+1} = w_k - A_k^{-1} g_k \tag{4}$$

When A_k is large, it is complex and time consuming to compute w_{k+1}, which makes Newton's method an alternative to the conjugate gradient methods for fast optimization (Ramesh et al. 2008).

Optimization of MLP structure

The learning process for MLP network trained using a BP algorithm consists of two iterative steps: forward computing of data stream and backward propagation of error signals. In this process, the synaptic weights are all adjusted in accordance with the error signals (Cai et al. 2009). In order to obtain a MLP model with the best performance, it is necessary to carry out an optimization of architectural and training parameters during the development of model. Optimization process, utilized in this study, was performed using a trial-and-error procedure that can be divided into three stages:

i. Determination of the optimum number of neurons in the hidden layer;
ii. Determination of the activation functions; and
iii. Determination of the BP algorithm which provides a model with the best performance.

Three different variations of BP algorithm (presented in "Back-propagation algorithm"), along with eight different activation functions shown in Table 4, were tested. ANN models were created using Neuroshell 2 software (Ward Systems Group Inc 2007).

Principal component analysis

The performance of ANNs depends mainly on data representation (Cherkassky and Lari-Najafi 1992). It is important to have uncorrelated inputs, since correlated data introduces confusion to the neural network during the learning process

Table 4 Activation functions tested in optimization stage

Activation function	Equation
Linear	$f(x)=x$
Logistic (sigmoid logistic)	$f(x)=\frac{1}{1+\exp(-x)}$
Symmetric logistic	$f(x)=\frac{2}{(1+\exp(-x))}-1$
Sine	$f(x)=\sin(x)$
Tanh (hyperbolic tangent)	$f(x)=\frac{\exp(x)-\exp(-x)}{\exp(x)+\exp(-x)}$
Tanh15 (hyperbolic tangent 1.5)	$f(x)=\tanh(1.5x)$
Gaussian	$f(x)=\exp(-x^2)$
Gaussian Complement	$f(x)=1-\exp(-x^2)$

(Tripathy 2010). PCA is a commonly used method for the transformation of inputs, which reduces the correlation between selected inputs (Alalawi et al. 2008). The applicability of PCA to the data sets used can be verified through the application of Kaiser-Meyer-Olkin (KMO) test of sampling adequacy and Bartlett's sphericity test. KMO values greater than 0.5 are acceptable (Kaiser 1974), while Bartlett's sphericity test gives the probability (p) that the correlation matrix is not an identity matrix, and therefore low p value indicates that the PCA is appropriate. Details for on the application of PCA are published elsewhere (Sousa et al. 2007; Noori et al. 2010).

Principal component regression

Multiple linear regression (MLR) is one of the most frequently used methods for prediction in many research fields, including air quality modeling (Agirre-Basurko et al. 2006; Sousa et al. 2007). MLR models are based on the determination of linear dependence of the output variable (Y) of the set of input variables (X_i), according to the equation:

$$Y = \alpha_0 + \sum_{i=1}^{k} \alpha_i X_i + \varepsilon_i \quad (5)$$

where, α_0 is a coefficient, α_i are the coefficients of the X_i independent variables (predictors) for a set of i observations, and ε_i is the residual error. In spite of its evident success in many applications, however, the regression approach can face serious difficulties when the independent variables are correlated with each other (Abdul-Wahab et al. 2005).

PCA is often used for the removal of multicollinearity of inputs before MLR is applied. After performing PCA on the original data, the obtained principal components (PCs) are used as independent variables in a regression equation. This approach, which combines MLR and PCA, is called PCR, and it establishes a relationship between the output variable and the selected PCs generated from the original input variables (Pires et al. 2008).

Performance indicators

In order to compare the performance of the different models, several statistical metrics presented in Table 5 were used. The index of model performance is a reformulation of the index of agreement, and it indicates the sum of the magnitudes of the differences between the model-predicted and observed deviations in connection with the observed mean value (Willmott et al. 2012). The Nash-Sutcliffe efficiency (NSE) is a normalized statistical approach that determines the relative magnitude of the residual variance ("noise") compared with the measured data variance ("information"). NSE indicates how well the plot of observed data, vs. the simulated data fits the 1:1 line (Moriasi et al. 2007). Mean absolute error (MAE) and root squared error (RMSE) measure residual errors, which give a global persepctive of the difference between the observed and modeled values (Sousa et al. 2007). Factor of 1.25 (FA1.25) represents the proportion of data for which the model results are "approximate" with the measured values, within the range of plus 25 % or minus 20 %.

Table 5 Performance indicators

Performance metric	Equation[a]
The index of model performance (d_r)	$d_r = 1 - \frac{\sum_{i=1}^{n} \lvert C_{p_i} - C_{o_i} \rvert}{2\sum_{i=1}^{n} \lvert C_{o_i} - \overline{C}_o \rvert}$
Nash-Sutcliffe efficiency (NSE)	$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (C_{o_i} - C_{p_i})^2}{\sum_{i=1}^{n} (C_{o_i} - \overline{C}_o)^2}$
The root squared error (RMSE)	$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n} (C_{o_i} - C_{p_i})^2}$
The mean absolute error (MAE)	$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} \lvert C_{o_i} - C_{p_i} \rvert$
RMSE-observation standard deviation ratio (RSR)	$\mathrm{RSR} = \frac{\mathrm{RMSE}}{\mathrm{Co}_{\mathrm{st.\ dev.}}}$
The percent of predictions within a factor of 1.25 of observed values (FA1.25)	$0.8 < \frac{C_p}{C_O} < 1.25$
The mean squared error (MSE)	$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (C_{o_i} - C_{p_i})^2$
The mean absolute percentage error (MAPE)	$\mathrm{MAPE} = 100\frac{1}{n}\sum \frac{\lvert C_o - C_l \rvert}{C_o}$

[a] C_p and C_o are the predicted and observed values. σ_o and σ_p are the standard deviation of observations and predictions. The over bar refers to the average of the values

Results and discussion

The optimization of MLP model

The optimization of MLP architecture begins with the determination of the optimal number of neurons in the hidden layer. This is performed using a trial-and-error approach: 16 different 4-H-1 architectures, where H was the number of hidden neurons that varied between 5 and 20, were tested and evaluated on the basis of mean square error (MSE) of the network (Table 5). The obtained MSE for MLP networks created during this optimization is presented in Fig. 1a. The optimum

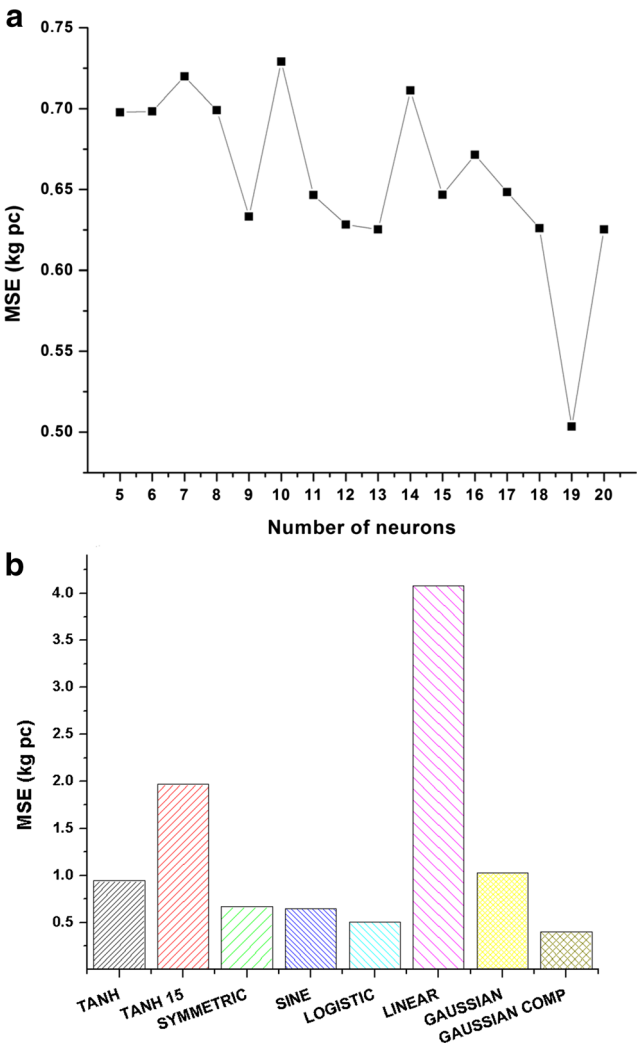

Fig. 1 The optimization of MLP architecture: **a** number of hidden neurons and **b** activation function

number of hidden neurons was found to be 19, and therefore the final MLP architecture was 4-19-1.

The second part of this optimization process is the determination of the best performing activation function among eight different variants which were considered (Table 4). Similarly as above, in order to find the optimal activation function, eight models with different activation functions were created and evaluated with respect to the MSE (Fig. 1b). As it can be seen, the MLP model with Gaussian Complement activation function demonstrated the best performance.

Finally, the optimization of architectural and training parameters was completed by determining the most suitable BP learning algorithm. In the present paper, three BP algorithms which are available in Neuroshell software were tested: Vanilla (learning rate gradient descent), Momentum (conjugate gradient), and TurboProp, the latter being a training method that operates much faster in "batch" mode and has an additional

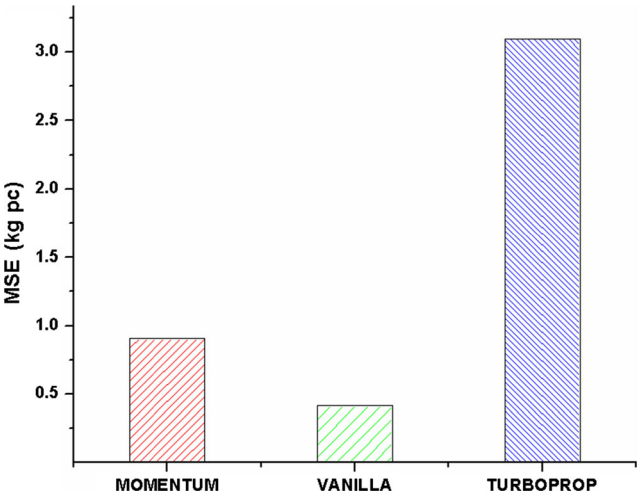

Fig. 2 MSE values for tested BP algorithms

advantage that it is not sensitive to learning rate and momentum (Ward Systems Group Inc 2008a). It can be concluded that the best performing MLP model is the one with 19 hidden neurons, Gaussian Complement activation function and Vanilla algorithm (Fig. 2).

PCA-MLP model

Additional improvement of a MLP model can be done by applying PCA to the input data. First, the applicability of PCA on available input data was tested under the terms of the KMO test and Bartlett's Test of Sphericity (Table 6). The obtained KMO value of 0.618 can be considered as acceptable, and Bartlett's Test of Sphericity also shows that PCA is appropriate ($p<0.0001$). Communality values, which range from 0 to 1 and indicate the amount of variability defined by each PC, are also presented in Table 6. It can be seen that for all variables, the extracted communality were greater than 0.8.

The Direct Oblimin method was used as a rotation method. The rotated factor loadings, structure matrix, eigenvalues, and respective variances as well as total variance explained by

Table 6 A summary of applied PCA: communalities extracted, KMO, and Bartlett's test values

Kaiser-Meyer-Olkin (KMO)	Approximate Chi-square	0.618
		225.911
Bartlett's test of sphericity	*df*	6
	Sig. (*p*)	<0.0001
Input	Communalities extracted	
FECT	0.905	
GDP	0.895	
CNF	1.000	
NC	0.999	

Table 7 Rotated factor loadings, structure matrix for PCA, eigenvalues, respective variances, as well as total variance explained PC

Input	Component					
	Rotated factor loadings			Structure matrix		
	PC1	PC1	PC3	PC1	PC2	PC3
FECT	0.970	−0.012	−0.041	0.951	0.073	0.392
GDP	0.921	0.014	0.049	0.945	0.107	0.464
CNF	0.000	1.000	−0.001	0.093	1.000	0.133
NC	0.002	−0.001	0.999	0.450	0.134	1.000
Eigenvalues				2.151	0.984	0.664
Variance per component (%)				53.76	24.61	16.60
Cumulative variance (%)				53.76	78.37	94.97

extracted PCs are shown in Table 7. Three PCs were extracted and the total variance explained by the three extracted PCs was 94.97 %.

Those extracted PCs were used as the input in the MLP model created using PCA-transformed inputs (PCA-MLP)-V model, which had architectural and training parameters which were the same as the MLP-V model created with original data. The obtained MSE for PCA-MLP-V was 0.45 kg NH_3 pc, while the comparison of actual and predicted NH_3 emissions using PCA-MLP-V model is presented in Fig. 3.

The validation of MLP models

The validation of created MLP models (MLP-V and PCA-MLP-V) was performed using a completely new dataset from

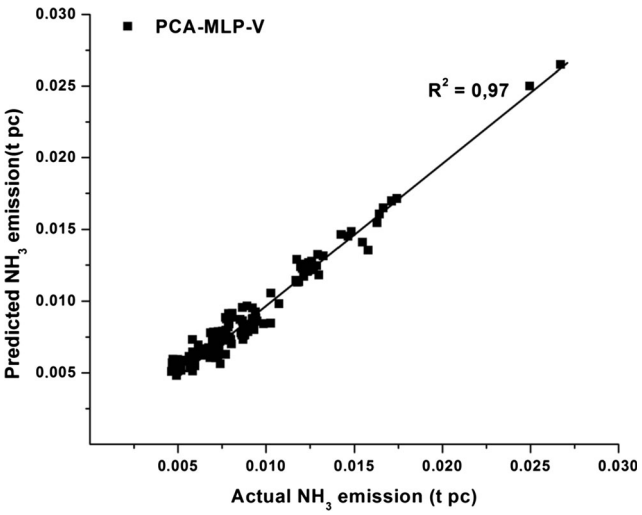

Fig. 3 Comparison of actual NH_3 emission and predicted NH_3 emission using PCA-MLP-V model

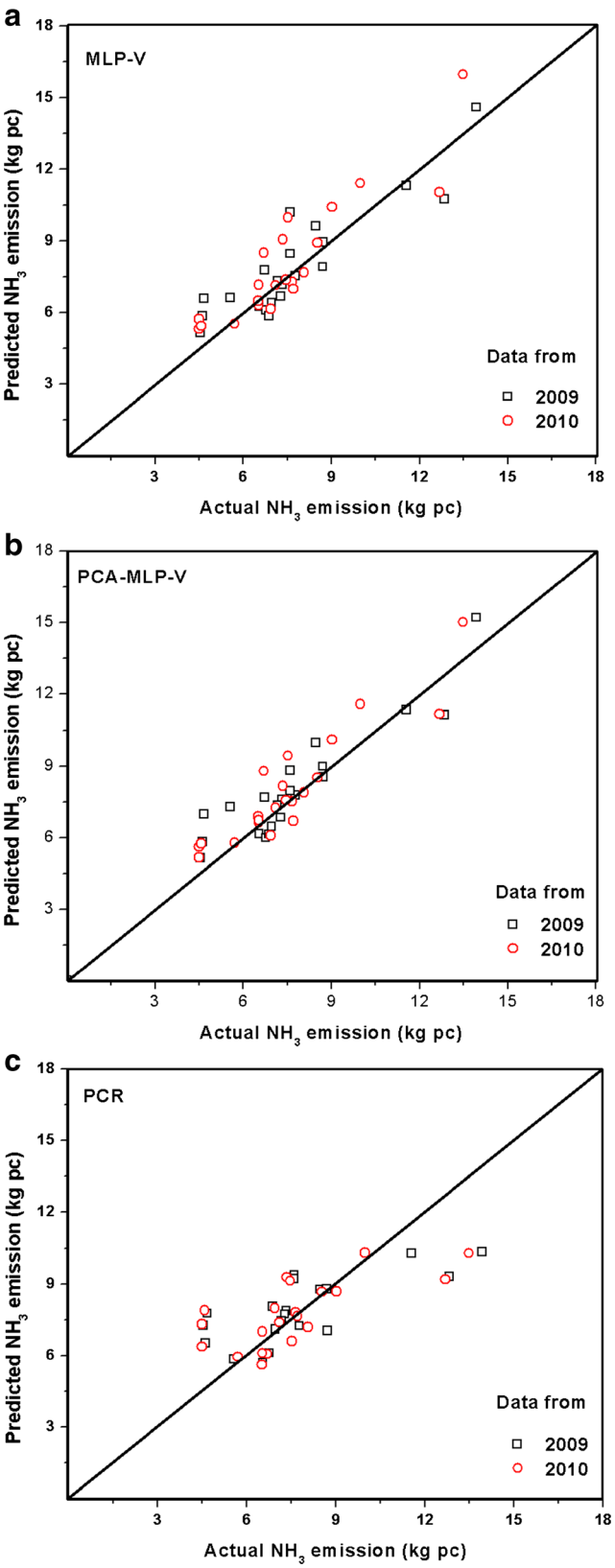

Fig. 4 Model validation results. Actual vs. predicted plot: **a** MLP-V, **b** PCA-MLP-V, and **c** PCR

the years 2009 and 2010 for all selected countries, the models being validated for a 1- and 2-year prediction.

For comparison with MLP models, PCR model (6) is created using PCs obtained for the PCA-MLP-V model.

$$NH_3 = 7.782 \cdot PC1 + 0.264 \cdot PC2 + 15.382 \cdot PC3 + 5.124 \quad (6)$$

Actual vs. predicted plot for MLP-V, PCA-MLP-V, and PCR models are presented in Fig. 4. The validation results have shown that both MLP models, MLP-V and PCA-MLP-V, produced significantly better prediction in comparison with linear PCR regarding all performance metrics (Table 8.).

Also, the results show that the PCA-MLP-V model performs slightly better than the MLP-V model, both for 1- and 2-year predictions. The application of PCA on the original inputs resulted in a decrease of RMSE and MAE values for about 10 %, while FA1.25 remains at a value of 86 %, which means that for the validation dataset the PCA-MLP-V model provided prediction with the relative error higher than 25 % in 14 % of cases. The values of mean absolute percentage error (MAPE) for the predictions made by the MLP-PCA-V and PCR models on the validation data (2009–2010) for the USA and EU countries are presented in Fig. 5. The MLP-PCA-V model provided better forecast results than the PCR model for the overall validation data (MAPE=11 %), for the USA and for most of the EU countries. Only in the cases of Bulgaria, Slovakia, and the UK, the MLP-PCA-V model has a prediction with MAPE value higher than the overall MAPE values of the PCR model of 18 %. In the case of the USA, the prediction of the MLP-PCA-V was with a relative error of 12 %, which is almost 2.5 times more accurate than the prediction obtained from the linear PCR model.

Regarding that the prediction of ammonia emissions with the proposed ANN approach has been performed using only four inputs, instead of a large number of emission factors and activity rates which are needed for inventory-based models, imposes this ANN approach as a viable alternative for the prediction of ammonia emissions at the national level. In particular, this ANN approach can be useful for developing countries that are also the largest emitters of pollutants into atmosphere, such as China, India and Russia, and which usually lack emission inventory data.

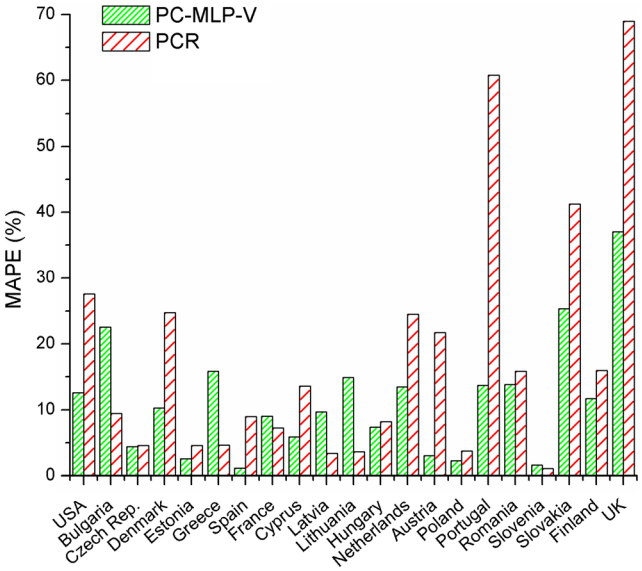

Fig. 5 Values of MAPE for validation data

Sensitivity analysis

Sensitivity analysis provides an understanding of the significance of model input variables. Sensitivity analysis in case of the BP networks can be carried out by determining the contribution factor for each input variable during the training process. The contribution factor is the sum of the absolute values of the weights leading from the particular variable (Ward Systems Group Inc 2008b). The number of cattle (NC) had the highest impact on the output (CF=14.05), followed by the GDP (CF=10.52). FECT (CF=5.67) and consumption of nitrogen fertilizers (CNF; CF=5.16) have a lower influence on the model output. A high contribution factor of NC only confirms the fact that agricultural sector dominates in the contribution of total NH3 emissions.

Conclusion

An ANN model which utilizes sustainability, economical, and agricultural indicators as inputs is proposed as an alternative approach for the prediction of annual NH_3 emissions at the

Table 8 Values of model performance indicators (validation stage)

Performance indicators	MLP-V model		PCA-MLP-V model		PCR model	
	2009	2009–2010	2009	2009–2010	2009	2009–2010
d_r	0.74	0.73	0.76	0.76	0.62	0.62
NSE	0.73	0.70	0.77	0.76	0.41	0.42
RMSE (kg pc)	1.07	1.13	1.00	1.02	1.69	1.66
MAE (kg pc)	0.86	0.90	0.80	0.80	1.29	1.24
RSR	0.46	0.48	0.43	0.43	0.72	0.71
FA 1.25 (%)	86	86	86	86	74	76

national level. Performance of the developed ANN models was validated by comparing predicted emissions with the published actual emission data for the USA and EU countries. The selected MLP architecture was optimized using a trial-and-error procedure, and PCA was applied to reduce mutual correlation between inputs. The obtained results demonstrated that the PCA-MLP provides a more accurate prediction in comparison with the MLP model based on original inputs.

Based on the results obtained, the developed ANN model can be considered as an alternative to conventional emission inventory based models for modeling ammonia emissions at the national level. The key advantage of the ANN model is that it uses a significantly lower number of input parameters that are all freely available for developed and even developing countries, and therefore can be applied in cases when the emission inventory approach cannot be applied because of a lack of data. Moreover, this approach can also be effectively implemented for the estimation of sectorial ammonia emissions.

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

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Statement of Human Rights and Statement on the Welfare of Animals This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abdul-Wahab SA, Bakheit CS, Al-Alawi SM (2005) Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. Environ Model Softw 20:1263–1271

Agirre-Basurko E, Ibarra-Berastegi G, Madariaga I (2006) Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area. Environ Model Softw 21:430–446

Alalawi S, Abdulwahab S, Bakheit C (2008) Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. Environ Model Softw 23:396–403

Amann M, Bertok I, Borken-Kleefeld J, Cofala J, Heyes C, Höglund-Isaksson L, Klimont Z, Nguyen B, Posch M, Rafaj P, Sandler R, Schöpp W, Wagner F, Winiwarter W (2011) Cost-effective control of air quality and greenhouse gases in Europe: modeling and policy applications. Environ Model Softw 26:1489–1501

Aneja VP, Blunden J, James K, Schlesinger WH, Knighton R, Gilliam W, Jennings G, Niyogi D, Cole S (2008) Ammonia assessment from agriculture: U.S. status and needs. J Environ Qual 37:515–520

Antanasijević DZ, Pocajt VV, Povrenović DS, Ristić MĐ, Perić-Grujić AA (2013) PM 10 emission forecasting using artificial neural networks and genetic algorithm input variable optimization. Sci Total Environ 443:511–519

Antanasijević DZ, Ristić MĐ, Perić-Grujić AA, Pocajt VV (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

Barrancos J, Briz S, Nolasco D, Melián G, Padilla G, Padrón E, Fernández I, Pérez N, Hernández PA (2013) A new method for estimating greenhouse gases and ammonia emissions from livestock buildings. Atmos Environ 74:10–17

Behera SN, Sharma M, Aneja VP, Balasubramanian R (2013) Ammonia in the atmosphere: a review on emission sources, atmospheric chemistry and deposition on terrestrial bodies. Environ Sci Pollut Res 20:8092–8131

Beusen AHW, Bouwman AF, Heuberger PSC, Van Drecht G, Van Der Hoek KW (2008) Bottom-up uncertainty estimates of global ammonia emissions from global agricultural production systems. Atmos Environ 42:6067–6077

Bishop CM (2006) Pattern recognition and machine learning, information. Springer, Singapore

Bolanča T, Cerjan-Stefanović Š, Ukić Š, Rogošić M, Luša M (2008) Application of different training methodologies for the development of a back propagation artificial neural network retention model in ion chromatography. J Chemom 22:106–113

Cai M, Yin Y, Xie M (2009) Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach. Transp Res Part D: Transp Environ 14:32–41

Cherkassky V, Lari-Najafi H (1992) Data representation for diagnostic neural networks. IEEE Expert 7:43–53

Cooter EJ, Bash JO, Benson V, Ran L (2012) Linking agricultural crop management and air quality models for regional to national-scale nitrogen assessments. Biogeosciences 9:4023–4035

Despagne F, Massart DL (1998) Neural networks in multivariate calibration. Analyst 123:157R–178R

Erisman JW, Sutton MA, Galloway J, Klimont Z, Winiwarter W (2008) How a century of ammonia synthesis changed the world. Nat Geosci 1:636–639

European Environment Agency (EEA) (2014) Ammonia (NH3) emissions. http://www.eea.europa.eu/data-and-maps/indicators/eea-32-ammonia-nh3-emissions-1. Accessed 20 May 2014

Eurostat (2014a) Final energy consumption by sector. http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tsdpc320&language=en. Accessed 18 May 2014

Eurostat (2014b) Statistics database. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database. Accessed 18 May 2014

Gardner M, Dorling S (1998) Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos Environ 32:2627–2636

Gilliland AB, Wyat Appel K, Pinder RW, Dennis RL (2006) Seasonal NH_3 emissions for the continental united states: inverse model estimation and evaluation. Atmos Environ 40:4986–4998

Gruber N, Galloway JN (2008) An Earth-system perspective of the global nitrogen cycle. Nature 451:293–296

Hájek P, Olej V (2012) Ozone prediction on the basis of neural networks, support vector regression and methods with uncertainty. Ecol Inform 12:31–42

Houlton BZ, Boyer E, Finzi AJ, Galloway J, Leach A, Liptzin D, Melillo J, Rosenstock TS, Sobota D, Townsend AR (2013) Intentional versus unintentional nitrogen use in the United States: trends, efficiency and implications. Biogeochemistry 114:11–23

Kaiser HF (1974) Index of factorial simplicity. Psychometrika 39:31–36

Karacan CÖ (2008) Modeling and prediction of ventilation methane emissions of U.S. longwall mines using supervised artificial neural networks. Int J Coal Geol 73:371–387

Li C, Salas W, Zhang R, Krauter C, Rotz A, Mitloehner F (2012) Manure-DNDC: a biogeochemical process model for quantifying

greenhouse gas and ammonia emissions from livestock manure systems. Nutr Cycl Agroecosysts 93:163–200

Lim Y, Moon Y-S, Kim T-W (2007) Artificial neural network approach for prediction of ammonia emission from field-applied manure and relative significance assessment of ammonia emission factors. Eur J Agron 26:425–434

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50:885–900

Noori R, Khakpour A, Omidvar B, Farokhnia A (2010) Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. Expert Syst Appl 37:5856–5862

Paulot F, Jacob DJ, Pinder RW, Bash JO, Travis K, Henze DK (2014) Ammonia emissions in the United States, Europe, and China derived by high-resolution inversion of ammonium wet deposition data: Interpretation with a new agricultural emissions inventory (MASAGE_ NH_3). J Geophys Res Atmos 119:4343–4364

Pinder RW, Pekney NJ, Davidson CI, Adams PJ (2004) A process-based model of ammonia emissions from dairy cows: improved temporal and spatial resolution. Atmos Environ 38:1357–1365

Pires JCM, Martins FG, Sousa SIV, Alvim-Ferraz MCM, Pereira MC (2008) Selection and validation of parameters in multiple linear and principal component regressions. Environ Model Softw 23: 50–55

Radojević D, Pocajt V, Popović I, Perić-Grujić A, Ristić M (2013) Forecasting of greenhouse gas emissions in Serbia using artificial neural networks. Energy Sources, Part A 35:733–740

Raimundo IM Jr, Narayanaswamy R (2001) Simultaneous determination of relative humidity and ammonia in air employing an optical fibre sensor and artificial neural network. Sensors Actuators B 74:60–68

Ramesh J, Vanathi PT, Gunavathi K (2008) Fault classification in phase-locked loops using back propagation neural networks. ETRI J 30: 546–554

Rotz CA, Montes F, Hafner SD, Heber AJ, Grant RH (2014) Ammonia emission model for whole farm evaluation of dairy production systems. J Environ Qual 43:1143–1158

Rumelhart DE, Hinton G, Williams R (1986) Learning representations by back-propagating errors. Nature 323:534–536

Russo A, Raischel F, Lind PG (2013) Air quality prediction using optimal neural networks with stochastic variables. Atmos Environ 79:822–830

Sousa S, Martins F, Alvimferraz M, Pereira M (2007) Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environ Model Softw 22:97–103

Sundvor I, Balaguer NC, Guerreiro C, Viana M, Querol X, Reche C, Amato F, Mellios G (2013) Road traffic's contribution to air quality in European cities ETC./ACM Technical Paper 2012/14.

Tripathy M (2010) Power transformer differential protection using neural network Principal Component Analysis and Radial Basis Function Neural Network. Simul Model Pract Theory 18:600–611

US Department of Agriculture (USDA) (2014) National Agricultural Statistics Service. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1017. Accessed 18 May 2014

US Energy Information Administration (USEIA) (2014) Total energy. http://www.eia.gov/totalenergy/data/annual/. Accessed 18 May 2014

US Environmental Protection Agency (EPA) (2014) National Emissions Inventory (NEI) air pollutant emissions trends data. http://www.epa.gov/ttnchie1/trends/. Accessed 18 May 2014

Vogl TP, Mangis JK, Zigler AK, Zink WT, Alkon DL DL (1988) Accelerating the convergence of the backpropagation method. Biol Cybern 59:256

Ward Systems Group Inc. (2007) Neuroshell 2 version 4.2. MD, USA

Ward Systems Group Inc. (2008) Backpropagation—weight updates. http://www.wardsystems.com/manuals/neuroshell2/index.html?backpropagation_weightupdates.htm

Ward Systems Group Inc. (2008) Contribution factors detail. http://www.wardsystems.com/manuals/neuroshell2/index.html?contributionfactorsdetail.htm

Webb J, Misselbrook TH (2004) A mass-flow model of ammonia emissions from UK livestock production. Atmos Environ 38:2163–2176

Willmott CJ, Robeson SM, Matsuura K (2012) A refined index of model performance. Int J Climatol 32:2088–2094

Wu W, Dandy GC, Maier HR (2014) Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. Environ Model Softw 54:108–127