International Journal of Environmental Science and Technology (2020) 17:4719–4732
https://doi.org/10.1007/s13762-020-02799-6

ORIGINAL PAPER

Applications of artificial neural networks and hybrid models for predicting CO_2 flux from soil to atmosphere

S. Altikat[1] · A. Gulbe[2] · H. K. Kucukerdem[1] · A. Altikat[3]

Received: 28 January 2020 / Revised: 22 May 2020 / Accepted: 2 June 2020 / Published online: 9 June 2020

Abstract
The goal of this research is to model the level of carbon dioxide flowing from soil to sky using various methods. The methods of multiple linear regression (MLR) and artificial neural networks (ANN) beside two different hybrid models were exploited to achieve this objective. These hybrid models were arranged as the prior two methods with principal component analysis (PCA). For the ANN, 36 different structures were used with different transfer (logsig–logsig, tansig–tansig, pureline–pureline, logsig–tansig, logsig–pureline and tansig–pureline)—learning functions (Levenberg–Marquardt and Gradient Descent with Momentum) and neuron numbers (10, 20 and 30). The manure norm, soil type, soil temperature, soil moisture content, soil depth, and photosynthetically active radiation values were taken into account as input parameters while CO_2 flux was output parameter. According to the research conducted, the best results were obtained from the ANN method. This method was followed by PCA + ANN, MLR and PCA + MLR methods. The R^2 value of the network established in the ANN method was determined as 0.98. In this ANN model, Levenberg–Marquardt and tansig–pureline with 30 neurons were used as transfer and learning functions, respectively. Besides, when principal components were used as input parameters, the lower R^2 values were obtained with both the MLR and ANN methods.

Keywords Artificial neural networks · Principal components · Linear regression · Saline soil · Soil moisture · Soil temperature

Introduction

There are a few main factors affecting soil CO_2 flux such as soil organic matter content, soil type, soil tillage and management systems, root respiration, etc. The decomposition of soil organic matter causes CO_2 flux (Kuzyakov 2002; Fender et al. 2013). Fertilization, especially N fertilization, accelerates CO_2 flux due to the effect of root development (Shao et al. 2013) and microbial activity (Yan et al. 2010; Fangueiro et al. 2008). Soil temperature and soil moisture affect soil CO_2 flux because of their direct impact on microbial activity (Risk et al. 2002; Rustad et al. 2001). Soil respiration amount increases with the increase in soil temperature (Kirschbaum 1995; William et al. 1994, Lou et al. 2003, Lu et al. 2008).

Various methods have been used while modeling of the CO_2 flux from soil to atmosphere. Assorted studies in the literature to model CO_2 fluctuation have applied various techniques (Oprea and Iliadis 2011; Ibarra-Berastegi et al. 2008; Huebnerova and Michalek 2014). Among these techniques, multiple linear regression and artificial neural networks have been mostly utilized (Huebnerova and Michalek 2014; Elangasinghe et al. 2014; Kurt and Oktay 2010; Banja et al. 2012).

ANN has been successfully utilized for modeling many complex systems (Droulia et al. 2009). This is an efficient method for modeling nonlinear systems. This method uses input and output parameters for prediction with different transfer-learning function combinations and neuron numbers (Franch and Panigrahi 1997). Besides, different neural

Editorial responsibility: Parveen Fatemeh Rupani.

S. Altikat
sefa.altikat@igdir.edu.tr

[1] Department of the Biosystems Engineering, Agriculture Faculty, Iğdır University, Iğdır, Turkey

[2] Department of the Computer Science, Vocational School of Technical Sciences, Iğdır University, Iğdır, Turkey

[3] Department of the Environmental Engineering, Engineering Faculty, Iğdır University, Iğdır, Turkey

network types have been used such as back-propagation neural network (Van Wijk and Bouten Verstraten 2002).

ANN is frequently applied in the studies on ecological modeling such as temperature and rainfall prediction (Somaratne et al. 2005; Zhuang et al. 2012; Wen et al. 2014; Li et al. 2017). Also, Papale and Valentini (2003), Hagen et al. (2006) and Song et al. (2014) stated that the ANN model is a very appropriate method for efficiently predicting soil respiration. In many studies, the ANN method has been used successfully to model gas emission in forest soils. For example, Van Wijk and Bouten Verstraten (2002) and Papale and Valentini. (2003) stated that CO_2 emissions successfully modeled in European forests using the ANN method.

MLR is another method that was chiefly used in modeling works (Hutchinson et al. 2000; Welles et al. 2001). Besides, MLR has been designated in most of the approaches to CO_2 modeling studies (Pedersen 2000). The model's performance in the researches using the MLR method has been evaluated considering R^2 values (Pedersen et al. 2001). Higher the R^2 value obtained from the study results approaches to 1, more accurate the model's acceptance (Hutchinson and Livingston 2001).

For the performance evaluation of this model, the R^2 values are taken into consideration. If the value of the R^2 approaches to 1, the efficiency of the model is considered as good. Bond-Lamberty and Thomson (2010) reported a linear model (R^2: 0.32) among the soil CO_2 flux, soil temperature and moisture. In this research, temperature and moisture values were used as inputs, CO_2 flux was used as output values. Similarly, Chen et al. (2013) obtained a R^2 of 0.40 from the linear model between the CO_2 flux and soil temperature—moisture contents.

The PCA method condenses the input parameters into a smaller set called principal components (Johnson and Wichern 2002). MLR and ANN are employed to model the levels of CO_2 flow from soil to the atmosphere. In addition to these methods, two different hybrid models were formed; one of the hybrid models was planned as PCA + MLR while the other was PCA + ANN. As for ANN, 36 different structures were used with different transfer—learning functions and number of neurons. The manure norm soil type, soil temperature, soil moisture content, soil depth and photosynthetically active radiation values were taken into account as input parameters while CO_2 flux was output parameter.

The level of CO_2 fluxed from the soil to the atmosphere is directly affected by factors such as soil type, fertilizer norm and application form of the fertilizer, soil temperature, soil moisture content and soil management practices. Continuous observation of CO_2, one of the most effective greenhouse gases in the atmosphere, is very important for a sustainable environmental approach. From this point of view, the ratio of CO_2 emitted from the soil to the atmosphere can be continuously monitored by modeling the level of CO_2. Artificial neural networks and hybrid models can determine the relationships between nonlinear changing factors and model these relationships with high accuracy.

The purpose of this research is to investigate the effects of different soil conditions on the fluxed CO_2 from soil to atmosphere and determine the best CO_2 flux model using artificial neural networks and hybrid models.

Table 1 Properties of soil examples

Soil properties	Normal soil	Saline soil
Soil texture	Clay-loam	Clay-loam
$CaCO_3$	6.53%	10.2%
EC	1228 µS cm^{-1}	5.48 µS cm^{-1}
pH	8	9.3

Materials and methods

Laboratory experiments

In this study, two different soil types (normal and saline), two different farmyard manure norms (2–4 t ha^{-1}) and two different manure application methods (surface and subsurface) were examined in the laboratory conditions for modeling CO_2 flux from soil to atmosphere.

Saline and normal type soil examples provided east of Iğdır pasture and west of Iğdır pasture, Turkey, respectively. In the east of Iğdır, pasture has saline soil properties. In this region, soils have salinity properties as a result of wrong field applications such as excess irrigation, conventional agriculture, etc. The properties of the soil used in laboratory experiments are given in Table 1.

The manure used in the experiments was applied with two different methods as surface and subsurface. Manure had been homogenously laid on the soil surface as surface application method. In the subsurface application, manure laid on the 10 cm soil depth and then mixed with a paddle. The chemical content of the farmyard manure is given in Table 2.

A flux-type temperature resistance was used in the laboratory experiments. The resistance is laid on the soil surface approximately 15 cm of soil depth. The electronic control unit was used for blocked temperature fluctuation. The automated ACE and soil CO_2exchange system were used

Table 2 Chemical content of the farmyard manure

Properties	Values
Organic matter	352 g kg^{-1}
pH	7.2
EC	3.4 dS m^{-1}
N	16 g kg^{-1}
P	8.2 g kg^{-1}
K	6.9 g kg^{-1}
Ca	65 g kg^{-1}
Mg	5.8 g kg^{-1}

for determining the CO_2 flux. The technical information of CO_2 exchange system is given in Table 3.

Before the experiments were started, the soil was saturated by water. After waiting for 2 days, the soil was heated from 20 to 50 °C degrees with grades of 0.5°. An electronic temperature control unit (ECU) with flexible temperature resistance was used for this purpose. After reaching the maximum temperature level, the temperature resistance and ECU system were deactivated until the soil temperature reached 20 °C. These processes were continued about 48 h for each factor.

The resistance equipped with an electronic control unit and the soil CO_2 exchange system is given in Fig. 1. Volumetric soil moisture percentage (%) and temperature (°C) were simultaneously measured via automated ACE and soil CO_2 exchange system sensors.

Dataset for CO_2 flux modeling

In the research, 27,713 data (7 parameters × 3959 observation) were used for CO_2 flux prediction model. These data were obtained by automated ACE and soil CO_2 exchange system during the 48 h for all of the factors.

The modeling with multiple linear regression

The MLR method and model architecture are given in Eq. 1 and Fig. 2, respectively. In the equation, Y is model's predicted value, X is contaminant concentration, a_i, i:0,…,n, is coefficient of regression.

$$Y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n \tag{1}$$

The MATLAB software was used for the MLR model. The input and output parameters for this model are given in Table 4.

The modeling with principal component analyses

The principal component analysis (PCA) was used to decrease the number of input parameters. These new

Table 3 Technical information of CO_2 exchange system

Technical specifications	Unit
Measurement of CO_2	Standard range: (Molar) approximately 40.0 mmols m^{-3}
Measurement of PAR	0–3000 μmols m^{-2} s^{-1} Silicon photocell
Measurement of soil temperature	6 selectable inputs for thermistors
Measurement of soil moisture	4 selectable inputs for industry-standard sensors
Flow control to chamber	200–5000 ml min^{-1} (137–3425 μmols s^{-1})
Flow control accuracy	± 3% of fsd
Chamber volume	Closed type 2.6 l/open type 1.0 l
Chamber diameter	230 mm

CO_2 flux device   Temperature resistance and ECU

Fig. 1 CO_2 flux, temperature resistance and electronic control unit

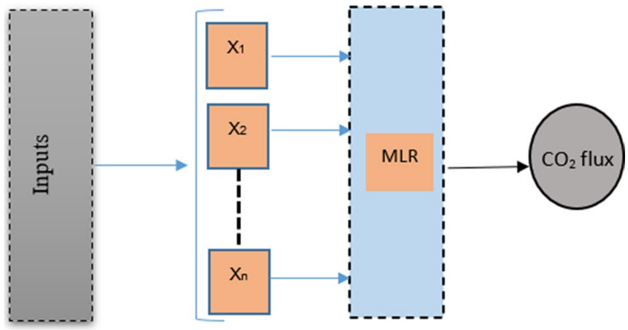

Fig. 2 Model architecture of MLR

input parameters were called principal components (PC-eigenvectors). To construct principal components MathWorks MATLAB was used. MATLAB's PCA function uses the singular value decomposition (SVD) algorithm by default and returns the percentage of the total variance explained by each principal component. In general, the smallest number of components explaining 80–99% of the total variance is chosen, where these values follow PCA best practices.

The modeling with artificial neural network (ANN)

Another model used in the research is artificial neural network (ANN). Artificial neural networks are frequently used in the modeling studies conducted between the variables which especially has nonlinear correlation. In this method, models are established with the aid of appropriate transfer and activation functions, number of neurons and learning algorithms considering the structural specifications of the problem (Gardner and Dorling 1998). In the research, the combinations of two learning functions, three transfer functions and three different neuron numbers were used in ANN structures to model CO_2 flux flowing from soil to air (Table 5). Artificial neural network architecture is given in Fig. 3.

Table 4 The input and output parameters

Input parameters	Output parameter
St: soil temperature (°C) Sm: soil moisture content (%) St: soil type Fn: fertilizer norm Sd: soil depth Pr: photosynthetically active radiation	CO_2 flux

Principal component analysis with multiple linear regression

In this method, for modeling CO_2 emission, PCs were accepted as input parameters and combined with the MLR method (Fig. 4). PCs were obtained from the principal component analyses.

Principal component analysis with artificial neural network

The PCs were used as input parameters in this method as in the PCA + MLR method. The same transfer—learning functions and neuron numbers used in the ANN method were used together with PCs for modeling CO_2 emission (Table 6). Figure 5 illustrates architecture of principal component analysis with artificial neural network.

Statistical analysis for the dataset

Analysis of variance (ANOVA) was used to assess the significance of each treatment on soil properties and CO_2 fluxes. Means were compared when the *F* test for treatment was significant at 5% level by using Duncan's multiple range tests.

Performance evaluation for hybrid models

Accuracies of models were confirmed via root mean-square error (which is also known as root mean-square deviation

Table 5 Functions and neurons numbers used in the ANN

Input parameters	ANN structures			Output parameter
	Learning functions	Transfer functions	Neurons	
St: soil temperature (°C) Sm: soil moisture content (%) St: soil type Fn: fertilizer norm Sd: soil depth Pr: photosynthetically active radiation	Levenberg–Marquardt (Trainlm) Gradient descent with momentum (Traingdm)	Logsig–logsig Tansig–tansig Pureline–pureline Logsig–tansig Logsig–pureline Tansig–pureline	10 20 30	CO_2 flux

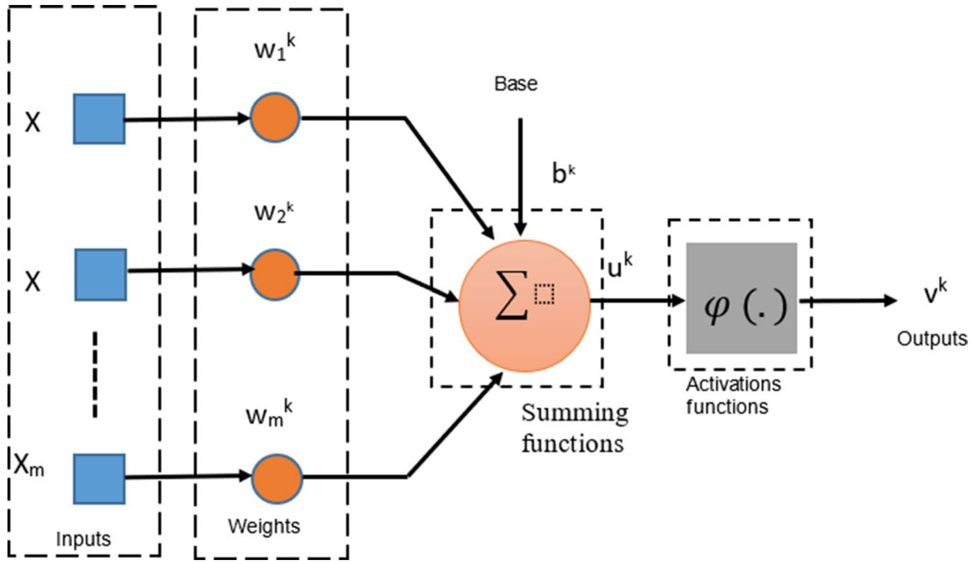

Fig. 3 Artificial neural network architecture

or RMSE), mean absolute error (MAE), and R^2 (which is also known as coefficient of determination or R^2). A model is evaluated as its accuracy is high when R^2 reaches to 1 and RMSE and MAE approaches to zero.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_{pi} - Y_{di}\right)^2} \tag{2}$$

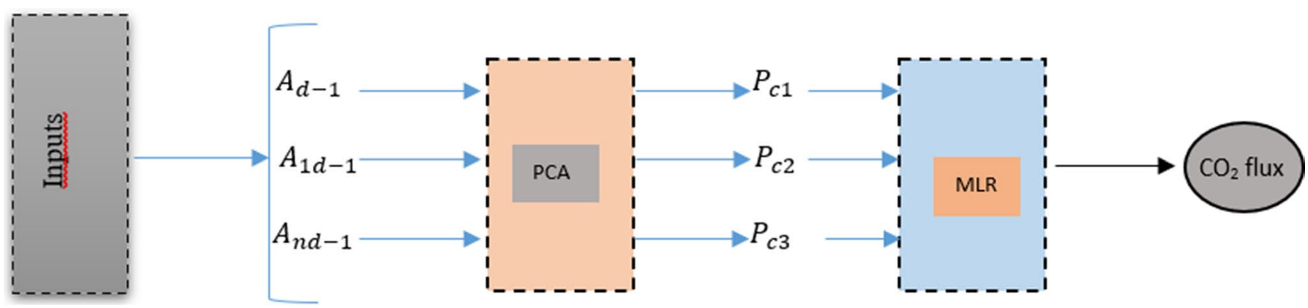

Fig. 4 The architecture of principle component analysis with multiple linear regression

Table 6 Functions and neurons numbers used in the PCA + ANN

Input parameters	ANN structures			Output parameter
	Learning functions	Transfer functions	Neurons	
PC_1	Levenberg–Marquardt (Trainlm)	Logsig–logsig	10 20 30	CO_2 flux
		Tansig–tansig		CO_2 flux
		Pureline–pureline		CO_2 flux
PC_2	Gradient descent with momentum (Traingdm)	Logsig–tansig		CO_2 flux
		Logsig–pureline		CO_2 flux
		Tansig–pureline		CO_2 flux

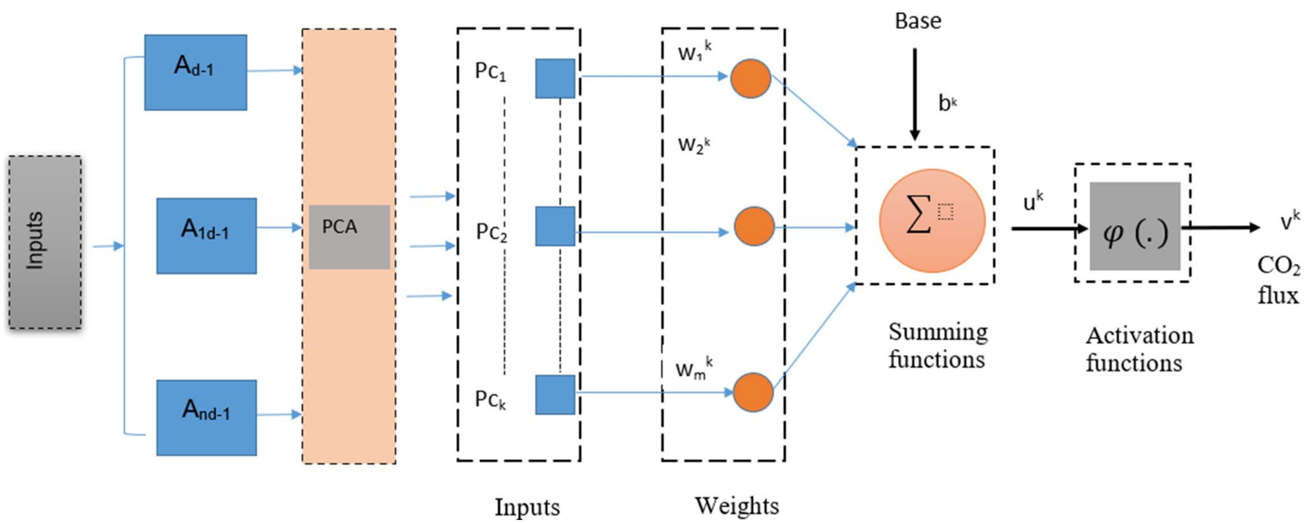

Fig. 5 The architecture of principal component analysis with artificial neural network

Table 7 The results of variance analysis for the dataset

	Factors	F	P
Main factors	Soil temperature	18.235	0.000**
	Soil type	3.782	0.050 *
	Fertilizer amount	9.108	0.006**
	The method of fertilizer applications	21.501	0.000**
Interactions	Temperature * soil type	0.269	0.926 ns
	Temperature * Manure norm	0.588	0.709 ns
	Temperature * soil depth	0.479	0.789 ns
Temperature			CO_2 flux
20–25 °C			1.173 c
25–30 °C			1.401 c
30–35 °C			2.350 c
35–40 °C			3.935 b
40–45 °C			4.705 b
45–50 °C			6.620 a
Soil type			CO_2 flux
Normal			3.758 a
Saline			2.971 b
Manure norm			CO_2 flux
2 t ha^{-1}			2.754 b
4 t ha^{-1}			3.975 a
Soil depth			CO_2 flux
Surface			4.303 a
Subsurface			2.426 b

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_{pi} - Y_{di}\right| \tag{3}$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}\left(Y_{pi} - Y_{di}\right)^2}{\sum_{i=1}^{n}\left(Y_{pi} - \bar{Y}\right)^2}\right) \tag{4}$$

In these equations, where n is the number of observations, Y_{pi} is the predicted value for observation i, Y_{di} is the real value from observation i, and $\bar{Y}$ is the average of the real value.

Results and discussion

The results of statistical analyses for the dataset

Soil CO_2 flux was affected by soil type, farmyard manure norm, manure application techniques and soil temperature statistically highly significant ($p < 0.001$), but this trend was not observed interaction values (Table 7).

At the initial temperature conditions (20–25 °C), CO_2 flux assigned as 1.173 µmol g cm^{-3}, CO_2 flux gradually raised according to higher soil temperature conditions. When the soil temperature had been reached the maximum level (45–50 °C), CO_2 flux from soil to atmosphere determined as 6.62 µmol g cm^{-3}. The CO_2 flux on the subsurface manure application was bigger than the surface manure application, approximately 50%. However,

Table 8 The statistical results for MLR analysis

	R^2	F	P	Estimated error variance
CO_2 flux	0.681	1408.081	0.000	2.197

the CO_2 flux increased with increasing manure norm. CO_2 flux determined as 2.754 and 3.975 µmol g cm^{-3} for 2 and 4 t ha^{-1} manure norm, respectively. When examined effects of soil type on the CO_2 flux, maximum CO_2 flux values were observed at the normal-type soil with 3.758 µmol g cm^{-3} and minimum values determined at the saline soil conditions with 2.971 µmol g cm^{-3}.

The results of multiple linear regression (MLR) modeling

In the research firstly, multiple linear regression models were used to estimate the CO_2 flux. For this purpose, soil temperature (St), soil moisture content (Sm), soil type (St), fertilizer norm (fn), soil depth (sd) and photosynthetically active radiation (PAR) were used as input parameters for prediction of CO_2 flux. Table 8 illustrates the statistical results of the MLR. Examining Table 8, it can be seen that R^2 and P values are 0.681 and 0.000, respectively. The equation of the MLR model and predicted—observed values are given in Eq. 5 and Fig. 6, respectively.

$$CO_2\ \text{flux} = -8.60 - 1.40x_1 + 1.1x_2 + 0.24x_3 + 0.03x_4 + 0.22x_5 9.045x_6 \quad (5)$$

In this equation, x_1: St, x_2: Sm, x_3: Sty, x_4: Fn, x_5: Sd, x_6: PAR.

Table 9 The eigenvalues of principal components analyses

PC_1 (%)	PC_2 (%)	PC_3 (%)	PC_4 (%)	PC_5 (%)	PC_6 (%)
82.39	15.77	1.37	0.29	0.17	0.001
98.17					

Table 10 The statistical results of PCs + MLR

	R^2	F	P	Estimated error variance
CO_2 flux	0.432	1504.418	0.000	3.912

The results of the principal component analysis (PCA) modeling

The principal component analysis (PCA) results showed that the first two principal components, PC_1 and PC_2, explained, respectively, 82.39 and 15.78% of the variance for all areas and jointly was responsible for more than 98.17% of the variance (Table 9). A similar result was found in a study by Panosso et al. (2011) on CO_2 fluxes, where the PCs together explained 70% of the variability of soil attributes (physical and chemical), with PC_1 explaining 52% and PC_2, 18%.

The result of multiple linear regression with principal component analyses (PCs + MLR) hybrid modeling

In this method, the PCs (PC_1 and PC_2) were used as input parameters to predict CO_2 flux. The results of the statistical analyses and the equation of this model are given in

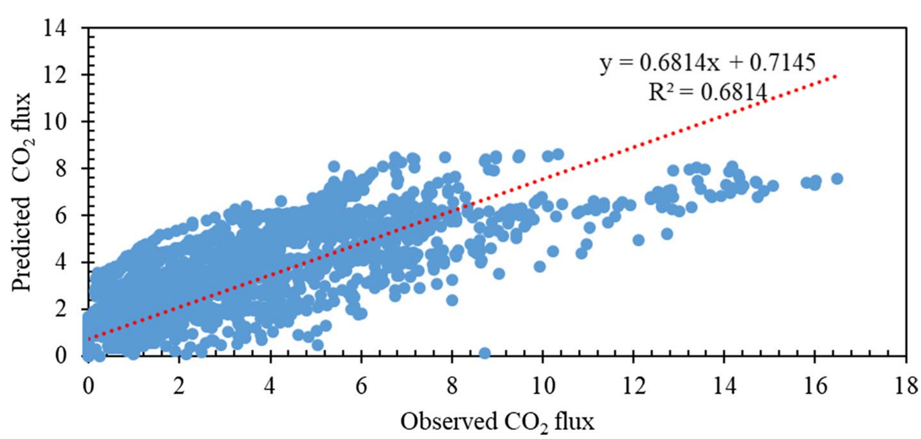

Fig. 6 Observed and predicted CO_2 flux in the multiple linear regression model

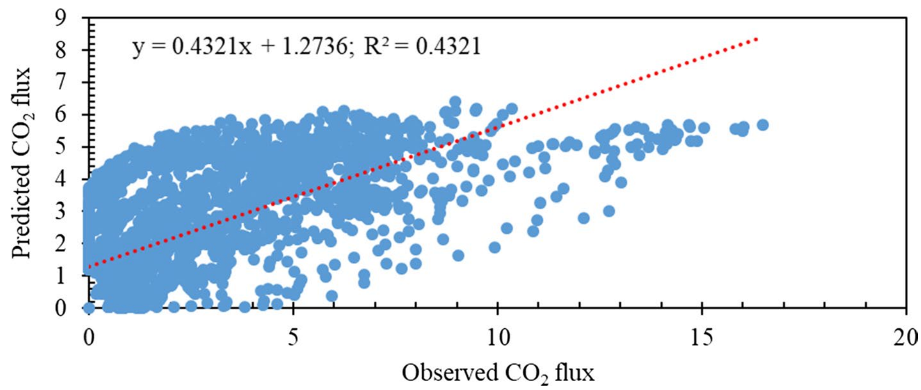

Fig. 7 Predicted and observed values of the PCs and MLR

Table 10 and Eq. 6, respectively. In this equation, x_1 and x_2 were expressed PC_1 and PC_2, respectively. Also, observed and predicted values can be seen in Fig. 7.

$$CO_2 \text{ flux} = -4.8 + 0.21x_1 + 0.05x_2 \tag{6}$$

The R^2 value was calculated as 0.432. This value is smaller than the R^2 of the MLR model. The MLR model used six input parameters such as soil temperature, soil moisture content, soil type, fertilizer norm, soil depth and photosynthetically active radiation for prediction of the CO_2 flux. However, this method used only two inputs parameters such as PC_1 and PC_2. According to this result, it can be said that better modeling will be done as the number of input parameters increase in the modeling of CO_2 emission.

The results of the artificial neural network (ANN) modeling

In the ANN, it was used 36 different neural structures with different learning—transfer functions with different neuron numbers. The statistical results of these structures are given in Table 11. Among these structures, the best results were obtained from the ANN 18 structure. This network model used Levenberg–Marquardt (Trainlm) learning function and Tansig–Pureline transfer function with 30 neurons.

In the ANN 18 structure, it can be seen that the highest R^2 and the lowest MAE values were calculated as 0.983 and 0.024, respectively. Also, the R values of test and validation were more than 0.99 (Fig. 8). Figure 9 illustrates the observed and predicted values of the ANN18 structure.

The results of the artificial neural network with principal components analysis (PCs and ANN) hybrid modeling

In this model, PCs were used as input parameters, and the 36 different ANN structures were examined for the CO_2 flux. Table 12 illustrates the statistical results of the PCs and the ANN model. The best-predicted results were obtained from the ANN16 structure. In this structure, the R^2 and MAE values were determined as 0.756 and 0.051, respectively. As can be seen in Table 12, the Levenberg–Marquardt (Trainlm) learning function and Logsig—Tansig transfer function with 30 neurons were used in the ANN 16 structure. Also, the R values training and validation were calculated as 0.86 and 0.87, respectively (Fig. 10).

The R^2 value of the PCs and ANN model was smaller than the ANN model. This result can be thought to be caused by the difference in input parameters. The PCA model has two input values such as PC_1 and PC_2, while the 6 input values (soil temperature, soil moisture content, soil type, fertilizer norm, soil depth and photosynthetically active radiation) in the ANN model have affected the model performance. Similar results were observed at the MLR and MLR with PCA models. Predicted and observed values of the CO_2 flux for Pcs and ANN model are given in Fig. 11.

Table 11 The statistical results of the ANN model

Model	Adaption learning function	Transfer function	Number of hidden neurons	RMSE	MAE	R^2
ANN $_1$	Trainlm	Logsig–logsig	10	0.795	0.749	0.010
ANN $_2$	Trainlm	Tansig–tansig	10	0.058	0.037	0.966
ANN 3	Trainlm	Pureline–pureline	10	0.180	0.123	0.681
ANN 4	Trainlm	Logsig–tansig	10	0.067	0.042	0.956
ANN 5	Trainlm	Logsig–pureline	10	0.061	0.039	0.963
ANN 6	Trainlm	Tansig–pureline	10	0.064	0.039	0.959
ANN 7	Trainlm	Logsig–logsig	20	0.791	0.741	0.293
ANN 8	Trainlm	Tansig–tansig	20	0.046	0.027	0.979
ANN 9	Trainlm	Pureline–pureline	20	0.180	0.123	0.681
ANN 10	Trainlm	Logsig–tansig	20	0.049	0.029	0.977
ANN 11	Trainlm	Logsig–pureline	20	0.050	0.030	0.975
ANN 12	Trainlm	Tansig–pureline	20	0.051	0.031	0.975
ANN 13	Trainlm	Logsig–logsig	30	0.791	0.741	0.294
ANN 14	Trainlm	Tansig–tansig	30	0.045	0.026	0.980
ANN 15	Trainlm	Pureline–pureline	30	0.180	0.123	0.681
ANN 16	Trainlm	Logsig–tansig	30	0.046	0.027	0.979
ANN 17	Trainlm	Logsig–pureline	30	0.042	0.025	0.983
ANN 18	Trainlm	Tansig–pureline	30	0.042	0.024	0.983
ANN 19	Traingdm	Logsig–logsig	10	0.795	0.749	0.000
ANN 20	Traingdm	Tansig–tansig	10	0.106	0.070	0.890
ANN 21	Traingdm	Pureline–pureline	10	0.180	0.123	0.681
ANN 22	Traingdm	Logsig–tansig	10	0.126	0.081	0.843
ANN 23	Traingdm	Logsig–pureline	10	0.117	0.077	0.864
ANN 24	Traingdm	Tansig–pureline	10	0.118	0.073	0.864
ANN 25	Traingdm	Logsig–logsig	20	0.795	0.749	0.257
ANN 26	Traingdm	Tansig–tansig	20	0.100	0.068	0.902
ANN 27	Traingdm	Pureline–pureline	20	0.180	0.123	0.681
ANN 28	Traingdm	Logsig–tansig	20	0.114	0.076	0.873
ANN 29	Traingdm	Logsig–pureline	20	0.134	0.090	0.823
ANN 30	Traingdm	Tansig–pureline	20	0.110	0.072	0.880
ANN 31	Traingdm	Logsig–logsig	30	0.795	0.749	0.053
ANN 32	Traingdm	Tansig–tansig	30	0.109	0.071	0.884
ANN 33	Traingdm	Pureline–pureline	30	0.180	0.122	0.681
ANN 34	Traingdm	Logsig–tansig	30	0.111	0.077	0.879
ANN 35	Traingdm	Logsig–pureline	30	0.121	0.076	0.856
ANN 36	Traingdm	Tansig–pureline	30	0.100	0.063	0.900

Conclusion

Among the methods conducted to model CO_2 flux, the ANN gave the best results. It is a good idea to visualize the data in the 2D plot using PCA to decrease 6 parameters to two principal components. However, PCA + MLR and PCA + ANN combinations resulted worse than MLR and ANN methods when they were utilized singularly. When PCs were constructed over 98% of the variance, it would be expected that MLR and ANN results should be close to PCA + MLR and PCA + ANN results, respectively. Regarding only R^2 values, MLR (0.681) differs from PCA + MLR (0.432) and ANN (0.983) from PCA + ANN (0.756). This much difference may be caused by our parameters being nonlinear. Further research may work on simultaneous-, progressive-, successive-, prioritized- (Liu

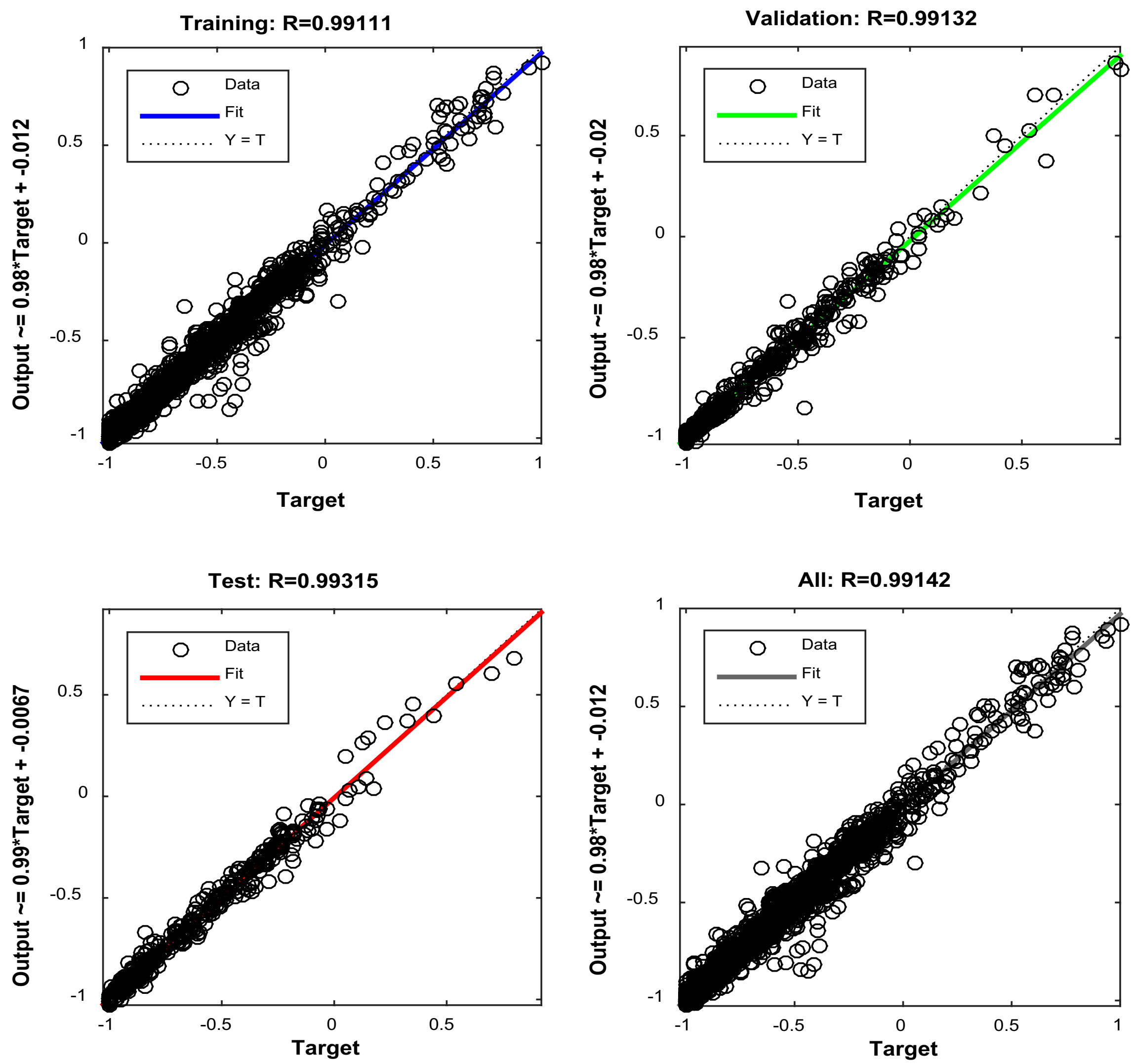

Fig. 8 Performance evaluation of the ANN18 structure for ANN model

Fig. 9 Predicted and observed CO_2 flux in the ANN_{18} structure for ANN model

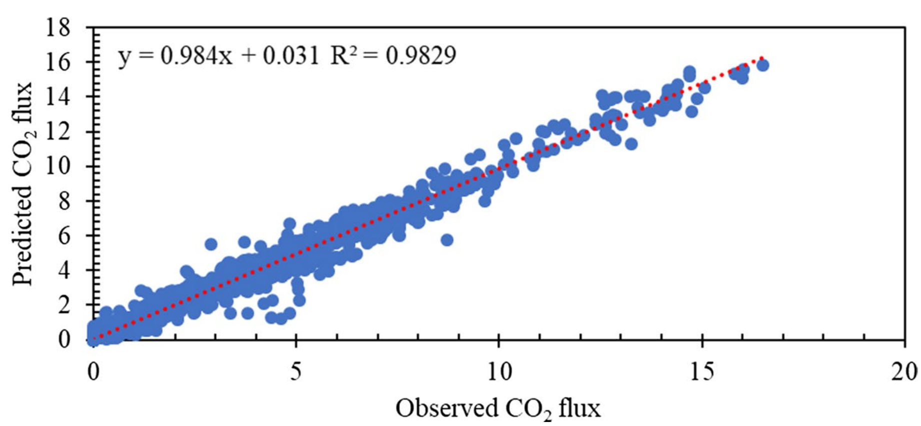

Table 12 The statistical results of the PCs and ANN model

Model	Adaption learning function	Transfer function	Number of hidden neurons	RMSE	MAE	R^2
ANN $_1$	Trainlm	Logsig–logsig	10	0.397	0.375	0.005
ANN $_2$	Trainlm	Tansig–tansig	10	0.088	0.055	0.692
ANN 3	Trainlm	Pureline–pureline	10	0.159	0.123	0.006
ANN 4	Trainlm	Logsig–tansig	10	0.089	0.056	0.684
ANN 5	Trainlm	Logsig–pureline	10	0.094	0.062	0.651
ANN 6	Trainlm	Tansig–pureline	10	0.091	0.059	0.673
ANN 7	Trainlm	Logsig–logsig	20	0.397	0.375	0.000
ANN 8	Trainlm	Tansig–tansig	20	0.091	0.061	0.675
ANN 9	Trainlm	Pureline–pureline	20	0.159	0.123	0.006
ANN 10	Trainlm	Logsig–tansig	20	0.084	0.052	0.724
ANN 11	Trainlm	Logsig–pureline	20	0.090	0.057	0.680
ANN 12	Trainlm	Tansig–pureline	20	0.093	0.060	0.657
ANN 13	Trainlm	Logsig–logsig	30	0.397	0.375	0.000
ANN 14	Trainlm	Tansig–tansig	30	0.082	0.052	0.732
ANN 15	Trainlm	Pureline–pureline	30	0.159	0.123	0.006
ANN 16	Trainlm	Logsig–tansig	30	0.079	0.051	0.756
ANN 17	Trainlm	Logsig–pureline	30	0.087	0.055	0.700
ANN 18	Trainlm	Tansig–pureline	30	0.082	0.051	0.732
ANN 19	Traingdm	Logsig–logsig	10	0.398	0.375	0.003
ANN 20	Traingdm	Tansig–tansig	10	0.153	0.118	0.081
ANN 21	Traingdm	Pureline–pureline	10	0.159	0.123	0.006
ANN 22	Traingdm	Logsig–tansig	10	0.157	0.122	0.037
ANN 23	Traingdm	Logsig–pureline	10	0.153	0.120	0.079
ANN 24	Traingdm	Tansig–pureline	10	0.151	0.117	0.105
ANN 25	Traingdm	Logsig–logsig	20	0.397	0.375	0.003
ANN 26	Traingdm	Tansig–tansig	20	0.148	0.113	0.134
ANN 27	Traingdm	Pureline–pureline	20	0.159	0.123	0.006
ANN 28	Traingdm	Logsig–tansig	20	0.150	0.115	0.114
ANN 29	Traingdm	Logsig–pureline	20	0.151	0.118	0.098
ANN 30	Traingdm	Tansig–pureline	20	0.142	0.101	0.209
ANN 31	Traingdm	Logsig–logsig	30	0.397	0.375	0.000
ANN 32	Traingdm	Tansig–tansig	30	0.143	0.109	0.194
ANN 33	Traingdm	Pureline–pureline	30	0.159	0.123	0.006
ANN 34	Traingdm	Logsig–tansig	30	0.208	0.133	0.012
ANN 35	Traingdm	Logsig–pureline	30	0.144	0.109	0.184
ANN 36	Traingdm	Tansig–pureline	30	0.144	0.106	0.184

and Chang 2007), independent-, sparse-, sparse independent- (Lee et al. 2016), parallel-, kernel- (Jiang and Yan 2018), local- or constrained (Aversano et al. 2019) PCA. ANN presented better models than MLR. Thus, CO_2 flux seems nonlinearly dependent on input parameters (manure norm, soil type, soil temperature, soil moisture content, soil depth, photosynthetically active radiation and maybe more). Nonlinear regression methods should be used instead of linear models.

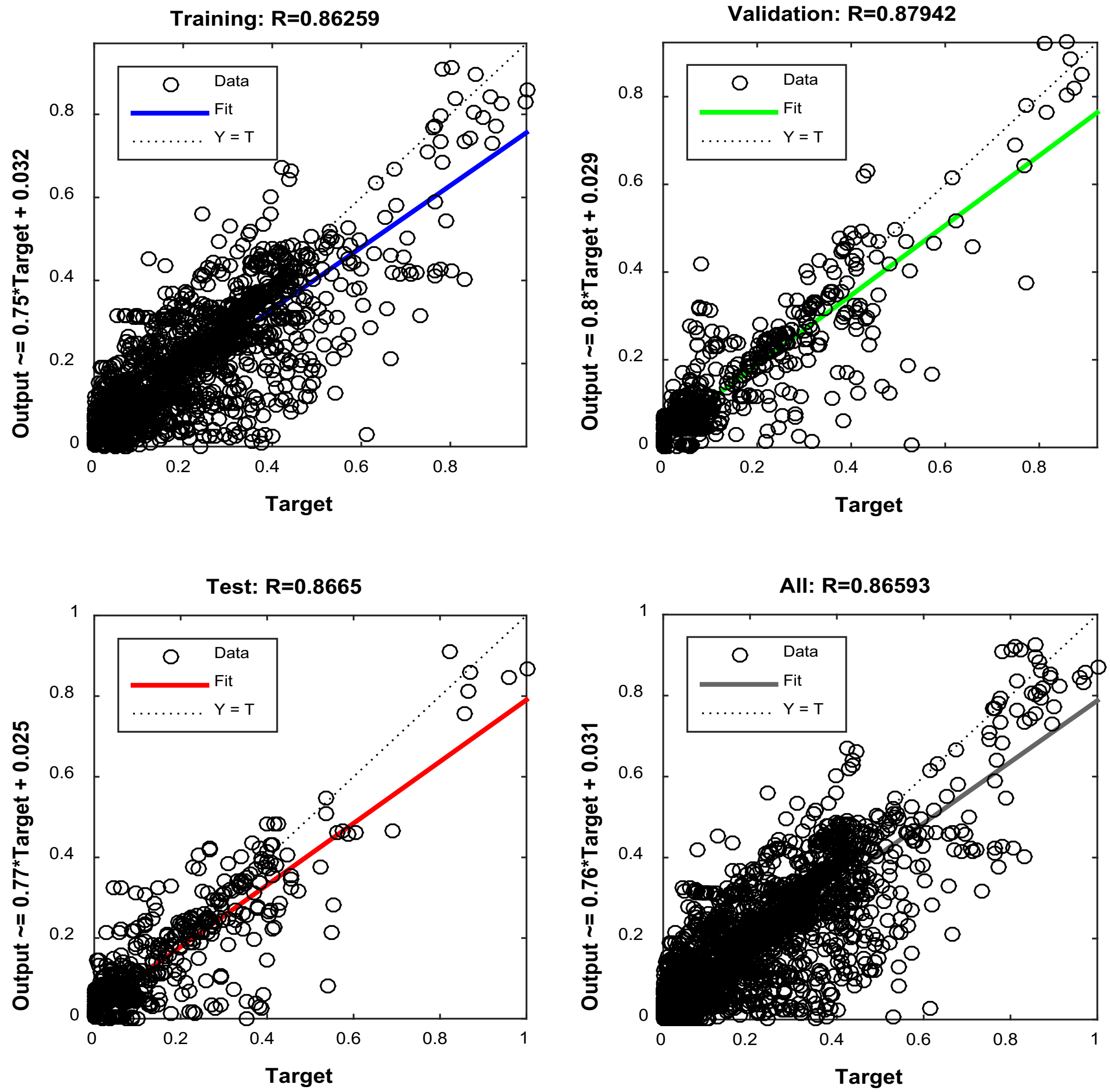

Fig. 10 Performance evaluation of the ANN16 structure for PCs and ANN model

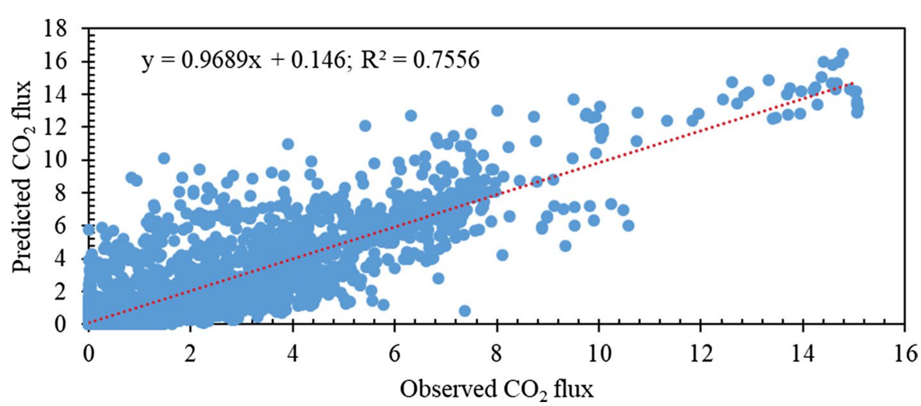

Fig. 11 Predicted and observed CO_2 flux in the ANN_{16} structure for PCs and ANN model

Acknowledgements This research was supported by the scientific research project unit of Iğdır University. The authors are thankful to the University of Igdir for providing the laboratory facilities.

Compliance with ethical standards

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

References

Aversano G, Bellemans A, Li Z, Coussement A, Gicquel O, Parente A (2019) Application of reduced-order models based on PCA & Kriging for the development of digital twins of reacting flow applications. Comput Chem Eng 121:422–441

Banja M, Papanastasiou DK, Poupkou A, Melas D (2012) Development of a short–term ozone prediction tool in Tirana area based on meteorological variables. Atmos Pollut Res 3:32–38

Bond-Lamberty B, Thomson A (2010) Temperature-associated increases in the global soil respiration record. Nature 464:579–582

Chen S, Huang Y, Xie W, Zou J, Lu Y, Hu Z (2013) A new estimate of global soil respiration from 1970 to 2008. Chin Sci Bull 58:4153–4160

Droulia F, Lykoudis S, Tsiros I, Alvertos N, Akylas E, Garofalakis I (2009) Ground temperature estimations using simplified analytical and semi-empirical approaches. Sol Energy 83:211–219

Elangasinghe MA, Singhal N, Dirks KN, Salmond JA (2014) Development of an ANN–based air pollution forecasting system with explicit knowledge through sensitivity analysis. Atmos Pollut Res 5:696–708

Fangueiro D, Senbayran M, Trindade H, Chadwick D (2008) Cattle slurry treatment by screw press separation and chemically enhanced settling: effect on greenhouse gas emissions after land spreading and grass yield. Bioresour Tech 99:7132–7142

Fender AC, Gansert D, Jungkunst HF, Fiedler S, Beyer F, Schutzenmeister K, Thiele B, Polle VK, Leuschner C (2013) Root-induced tree species effects on the source/sink strength for greenhouse gases (CH_4, N_2O and CO_2) of a temperate deciduous forest soil. Soil Biol Biochem 57:587–597

Franch LJ, Panigrahi S (1997) Artificial neural network models of wheat leaf wetness. Agric For Meteorol 88:57–65

Gardner MW, Dorling SR (1998) Artificial neural networks (the multilayer perceptron): a review of applications in the atmospheric science. Atmos Environ 32:2627–2636

Hagen SC, Braswell BH, Linder E, Frolking S, Richardson AD, Hollinger DY (2006) Statistical uncertainty of eddy flux-based estimates of gross ecosystem carbon exchange at Howland Forest, Maine. J Geophys Res Atmos 111:1–12

Huebnerova Z, Michalek J (2014) Analysis of daily average PM_{10} predictions by generalized linear models in Brno, Czech Republic. Atmos Pollut Res 5:471–476

Hutchinson GL, Livingston GP (2001) Vents and seals in nonsteady state chambers used for measuring gas exchange between soil and the atmosphere. Eur J Soil Sci 52:675–682

Hutchinson GL, Livingston GP, Healy RW, Striegl RG (2000) Chamber measurement of surface-atmosphere trace gas exchange: numerical evaluation of dependence on soil, interfacial layer, and source/sink properties. J Geophys Res 105(D7):8865–8875

Ibarra-Berastegi G, Elias A, Barona A, Saenz J, Ezcurra A, de Argandona JD (2008) From diagnosis to prognosis for forecasting air pollution using neural networks: air pollution monitoring in Bilbao. Environ Model Softw 23:622–637

Jiang Q, Yan X (2018) Parallel PCA–KPCA for nonlinear process monitoring. Control Eng Pract 80:17–25

Johnson RA, Wichern DW (2002) Applied multivariate statistical analysis, 5th edn. Prentice Hall, Upper Saddle River

Kirschbaum MUF (1995) The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. Soil Biol Biochem 27:753–760

Kurt A, Oktay AB (2010) Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. Expert Syst Appl 37:7986–7992

Kuzyakov Y (2002) Review: factors affecting rhizosphere priming effects. J Plant Nutr Soil Sci 165:382–396

Lee LC, Liong CY, Osman K, Jemain AA (2016) Comparison of several variants of principal component analysis (PCA) on forensic analysis of paper based on IR spectrum. In: AIP conference proceedings, AIP Publishing, vol 1750, pp 1–6. https://aip.scitation.org/doi/pdf/10.1063/1.4954617?class=pdf. Accessed 22 Dec 2019

Li P, Peng C, Wang M, Li W, Zhao P, Wang K, Yang Y, Zhu Q (2017) Quantification of the response of global terrestrial net primary production to multifactor global change. Ecol Indicat 76:245–255

Liu WM, Chang CI (2007) Variants of principal components analysis. IEEE, 1-4244-1212-9/07. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4422989. Accessed 22 Dec 2019

Lou YS, Li Z, Zhang TL (2003) Soil CO_2 flux in relation to dissolved organic carbon, soil temperature and moisture in a subtropical arable soil of China. J Environ Sci 15(5):715–720

Lu X, Cheng G, Xiao F, Fan J (2008) Modeling effects of temperature and precipitation on carbon characteristics and GHGs emissions in Abies fabric forest of subalpine. J Environ Sci 20(3):339–346

Oprea M, Iliadis L (2011) An artificial intelligence–based environment quality analysis system. In: Engineering applications of neural networks, Pt I, vol 363, pp 499–508. https://link.springer.com/content/pdf/10.1007%2F978-3-642-23957-1_55.pdf. Accessed 22 Dec 2019

Panosso AR, Marques Júnior J, Milor DMBP, Ferraudo AS, Barbieri M, Pereira GT, La Scala N (2011) Soil CO_2 emission and its relation to soil properties in sugarcane areas under Slash-and-burn and Green harvest. Soil Tillage Res 111:190–196

Papale D, Valentini R (2003) A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biol 9:525–535

Pedersen AR (2000) Estimating the nitrous oxide emission rate from the soil surface by means of a diffusion model. Scand J Stat 27:385–403

Pedersen AR, Petersen SO, Vinther FP (2001) Stochastic diffusion model fro estimating trace gas emissions with static chambers. Soil Sci Soc Am J 65:49–58

Risk D, Kellman L, Beltrami H (2002) Carbon dioxide in soil profiles: production and temperature dependence. Geophys Res Lett 29(6):11–14

Rustad LE, Campbell JL, Marion GM (2001) A metaanalysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. Oecologia 126:543–562

Shao C, Chen J, Li L (2013) Grazing alters the biophysical regulation of carbon fluxes in a desert steppe. Environ Res Lett 8(2):025012

Somaratne S, Seneviratne G, Coomaraswamy U (2005) Prediction of soil organic carbon across different land-use patterns: a neural network approach. Soil Sci Soc Am J 69:1580–1589

Song X, Peng C, Zhao Z, Zheng Z, Guo B, Wang W, Jiang H, Zhu Q (2014) Quantification of soil respiration in forest ecosystems across China. Atmos Environ 9:546–551

Van Wijk MT, Bouten Verstraten JM (2002) Comparison of different modeling strategies for simulating gas exchange of a Douglas-fir forest. Ecol Model 158:63–81

Welles JM, Demetriades-Shah TH, McDermitt DK (2001) Considerations for measuring ground CO_2 effluxes with chambers. Chem Geol 177:3–13

Wen X, Zhao Z, Deng X, Xiang W, Tian D, Yan W, Zhou X, Peng C (2014) Applying an artificial neural network to simulate and predict Chinese fir (*Cunninghamia lanceolata*) plantation carbon flux in subtropical China. Ecol Model 294:19–26

William T, Peterjon J, Melilio M, Paul A, Steudler A, Kathleen M (1994) Response of trace gas fluxes and N availability to experimentally elevated soil temperatures. Ecol Appl 4(3):617–625

Yan L, Chen S, Huang J, Lin G (2010) Differential responses of auto- and heterotrophic soil respiration to water and nitrogen addition in a semiarid temperate steppe. Glob Chang Biol 16:2345–2357

Zhuang Q, Lu Y, Chen M (2012) An inventory of global N_2O emissions from the soils of natural terrestrial ecosystems. Atmos Environ 47:66–75