


# Forecasting the spatiotemporal variability of soil CO<sub>2</sub> emissions in sugarcane areas in southeastern Brazil using artificial neural networks

Luciana P. S. Freitas · Mara L. M. Lopes · Leonardo B Carvalho · Alan R. Panosso ·  
Newton La Scala Júnior · Ricardo L. B. Freitas · Carlos R. Minussi ·  
Anna D. P. Lotufo 

Received: 14 August 2018 / Accepted: 12 November 2018  
© Springer Nature Switzerland AG 2018

**Abstract** Carbon dioxide (CO<sub>2</sub>) is considered one of the main greenhouse effect gases and contributes significantly to global climate change. In Brazil, the agricultural areas offer an opportunity to mitigate this effect, especially with the sugarcane crop, since, depending on the management system, sugarcane stores large amounts of carbon, thereby removing it from the atmosphere. The CO<sub>2</sub> production in soil and its transport to the atmosphere are the results of biochemical processes such as the decomposition of organic matter and roots and the respiration of soil organisms, a phenomenon called soil CO<sub>2</sub> emissions (FCO<sub>2</sub>). The objective of the study was to investigate the use of neural networks with backpropagation algorithm to predict the spatial patterns

of soil CO<sub>2</sub> emission during short periods in sugarcane areas. FCO<sub>2</sub> values were collected in three commercial crop areas in the São Paulo state, southeastern Brazil, registered through the LI-8100 system during the years 2008 (Motuca), 2010 (Guariba city), and 2012 (Pradópolis), in the period after the mechanical harvesting (green cane). A neural network multilayer perceptron with a backpropagation algorithm was applied to estimate the FCO<sub>2</sub> in 2012, using data from 2008 and 2010 as training for the neural network. The neural network initially presented a mean absolute percentage error (MAPE) of 18.3852 and a coefficient of determination ( $R^2$ ) of 0.9188. Data obtained from the observed and estimated values of FCO<sub>2</sub> present moder-

L. P. S. Freitas · C. R. Minussi · A. D. P. Lotufo (✉)  
Department of Electrical Engineering, UNESP - São Paulo State  
University, Campus of Ilha Solteira, Ilha Solteira, SP, Brazil  
e-mail: annadiva@ieee.org

L. P. S. Freitas  
e-mail: melscarin@gmail.com

C. R. Minussi  
e-mail: crminussi@gmail.com

M. L. M. Lopes · A. R. Panosso  
Department of Mathematics, UNESP - São Paulo State University,  
Campus of Ilha Solteira, Ilha Solteira, SP, Brazil

M. L. M. Lopes  
e-mail: mara.lopes@unesp.br

A. R. Panosso  
e-mail: alan.panosso@unesp.br

L. B. Carvalho  
Department of Plant Protection, UNESP - São Paulo State  
University, Campus of Jaboticabal,  
Jaboticabal, SP, Brazil  
e-mail: lbcarvalho@fcav.unesp.br

N. La Scala Júnior  
Department of Exact Sciences, UNESP - São Paulo State  
University, Campus of Jaboticabal,  
Jaboticabal, SP, Brazil  
e-mail: lascala@fcav.unesp.br

R. L. B. Freitas  
Department of Electrical Engineering,  
Western Parana State University,  
Foz do Iguaçu, PR, Brazil

ate spatial dependence, and it is observed from the maps of the spatial pattern of the CO<sub>2</sub> flow that the results from the neural network show considerable similarity to the observed data. The model results identify the higher and lower characteristics in sample points of CO<sub>2</sub> emissions and produce an overestimation of the range of spatial dependence (0.45 m) and an underestimation of the interpolated values in the field ( $R^2 = 0.80$ ; MAPE = 12.0591), when compared to the actual soil CO<sub>2</sub> emission values. Therefore, the results indicate that the artificial neural network provides reliable estimates for the evaluation of FCO<sub>2</sub> from data of the soil's physical and chemical attributes and describes the spatial variability of FCO<sub>2</sub> in sugarcane fields, thereby contributing to the reduction of uncertainties associated with FCO<sub>2</sub> accountings in these areas.

**Keywords** Neural techniques · Soil respiration · Green harvest

## Introduction

The human influence on the climate system is clear and recent levels anthropogenic greenhouse gas emissions are the highest ever (IPCC 2014). In addition, policies governing agricultural practices and the conservation and management of forests are most effective when they involve mitigation and adaptation, thus providing opportunities to reduce impacts and increase risk management. The increase of anthropogenic emissions of greenhouse gases (GHGs) is the dominant cause of global warming. Brazil, despite not having binding targets to reduce CO<sub>2</sub>, made a voluntary commitment to reduce their emissions by the year 2020, being a signatory to several multilateral agreements to the global effort to mitigate GHGs. Unlike other countries, whose burning of fossil fuels is the main source of GHG emissions, Brazil's emissions are mainly from changes in land use and deforestation.

Globally, agriculture is responsible for significant amounts of anthropogenic emissions of the greenhouse gases, namely, CO<sub>2</sub>, methane gas (CH<sub>4</sub>), and nitrous oxide (N<sub>2</sub>O), to the atmosphere, contributing 11%, 47%, and 58% of total greenhouse gas emissions, respectively (IPCC 2007). In Brazil, the agricultural share of CO<sub>2</sub> emissions represents 75% of the total anthropogenic emissions, due of the use of and changes in the land associated with such activity

(Cerri et al. 2007). However, depending on the use and management of agricultural land, it can act as a source or a sink of atmospheric carbon (Ussiri and Lal 2009). According to Rayment and Jarvis (2000), the carbon balance of the ecosystem is represented by the quantity of carbon absorbed by photosynthesis and the amount of carbon lost in the soil respiration process. Therefore, small increases in respiration rates may be sufficient to change an ecosystem from a sink to a source of carbon to the atmosphere. Frequently, soil emissions associated with the loss of soil carbon via CO<sub>2</sub> emissions in agricultural areas are not considered because of the emission's large timeline variation and because it is a phenomenon resulting from a complex interaction of physical, chemical, and biological attributes and climatic variables (Epron et al. 2006; Figueiredo and La Scala Júnior 2011; Teixeira et al. 2012; Moitinho et al. 2013).

In the agricultural context, Brazil stands out as the world's largest producer of sugarcane (*Saccharum* spp.) with a production of 625.96 million tons for the 2018/2019 crop (CONAB 2018). The harvested area was estimated at 8.61 million hectares, distributed in all producing states, with the State of São Paulo being the largest producer, accounting for 54% of the processed sugarcane. Thus, estimates of soil CO<sub>2</sub> emission in sugarcane areas are important to the reduction of spatiotemporal uncertainties associated with the carbon balance calculation process for the country. In southeastern Brazil, changes in the management of sugarcane crop practices are important because large areas are being converted from a burnt cane management system to a green harvest system.

The addition of larger mass organic residues in the green harvest system (green cane) causes a larger interaction between organic and mineral fractions of soil as well as the protection of the ground surface against the impact of rain drops, thus reducing the erosion process. The return of crop residues to the soil surface favors the accumulation of organic matter and a reduction in gas emissions compared to the burned cane system (Razafimbelo et al. 2006; Cerri et al. 2007; Luca et al. 2008). The use of new methodologies to forecast the spatiotemporal patterns of FCO<sub>2</sub> is necessary to improve the quality of current estimates. Thus, the use of artificial neural networks is important in this task.

## Neural networks

Artificial neural networks (ANNs) are an excellent alternative for modeling nonlinear systems that are too complex to be described by analytical methods, and ANNs are able to recognize patterns and perform diagnoses and predictions (Haykin 1999). ANNs are defined as adaptive communication networks linking a cost function to a desired output (Kartalopoulos 1996). ANNs can also be defined as a network that represents a dynamic process that can be modeled by a set of coupled nonlinear differential and/or algebraic equations (Kartalopoulos 1996). However, ANNs need to learn the task for which they were designed, and this is associated with the weights that connect the network. Therefore, knowledge is acquired by the network through a learning process and connection strengths between neurons (synaptic weights), which are then used to store the acquired knowledge (Haykin 1999).

Neural networks may have a single layer or multiple layers. The neural network used in this work is a multilayer neural network called the multilayer perceptron (MLP) trained by the algorithm backpropagation (BP). MLP is a type of feedforward neural network which is composed of many simple perceptron in a hierarchical structure forming a feedforward topology with one or more layers (hidden layers) between input and output layers (Kartalopoulos 1996; Haykin 1999).

The ability to learn through example and generalize the learned information make ANNs an interesting proposition for solutions in various areas of knowledge, such as engineering, medicine, meteorology, and agricultural science. One of the main neural network characteristics is the ability to learn and improve its performance. The network training introduces an input data set that produces an estimated or at least a consistent output, according to Wasserman (1989). Next, the calculated output is compared with a desired output, and thus an error is produced. This error is minimized in backward sense using the gradient descent algorithm until the minimum one is acquired (Haykin 1999). The weights are recalculated in each iteration, and this is the training process known as BP, introduced by Werbos (1974) and later improved by Widrow and Lehr (1990).

The Levenberg-Marquardt training algorithm is similar to backpropagation but the gradient descent minimization algorithm is replaced by a second-order Newton method (second derivative) (Liu 2010).

GRNN (general regression neural network) neural network (Spetch 1991) has a different concept when compared to a MLP, once it is based on a probability density function. It has four layers, one being the input, two hidden, and one output. The quantity of neurons in each layer depends on the quantity of inputs, outputs, and the training dataset. This neural network is trained in one presentation of the training patterns.

The MLP has been successfully used to solve problems in several areas involving high degrees of nonlinearity (Haykin 1999). Therefore, the use of an ANN is currently very significant. Thus, the works described below emphasize applications in agricultural.

Lotufo et al. (2007) introduced a procedure for the transient stability analysis and preventive control of electric power systems, and this procedure was formulated with a multilayer feedforward neural network. The neural network training was performed using the backpropagation algorithm with a fuzzy controller, and the most accurate results were obtained with this procedure compared to the traditional backpropagation algorithm. The work of Moretti et al. (2016) with the modulus of elasticity and compressive strength of concrete specimens used artificial neural networks with the backpropagation technique to predict the values of the elasticity modulus and resistance to concrete compression. Lentzsch et al. (2005) used a regression model and an artificial neural network with a set of 89 data points to evaluate the soil microbial biomass (SMB) of 30 sites in Germany, using a multilayer feedforward neural network built and tested with the Levenberg-Marquard algorithm for training. This model used a configuration with two or four hidden neurons and one output neuron. The results showed that both the linear full-factorial regression model and the neural network model are promising tools for predicting the SMB.

Merdun (2011) used a self-organizing neural network (SOM of Kohonen) to analyze the effects of the physical properties of soil on chemical/hydraulic processes and to diagnose the interrelationships of the multivariable soil data in the vadose zone. His conclusion shows that this neural network is effective for analyzing and diagnosing the dynamics of soil. Kim and Gilley (2008) used a multilayer perceptron neural network trained by backpropagation to estimate dissolved P (DP) and NH<sub>4</sub>-N concentrations in the soil erosion runoff from a land application site near Lincoln, NE, USA. The results showed that these estimations are

reliable. Song et al. (2014) analyzed the FCO<sub>2</sub> in Chinese forests using soil attributes and meteorological variables. They observed that a MLP neural network with three layers, trained with backpropagation algorithm and the gradient descent method, was an effective approach to predict FCO<sub>2</sub> in all Chinese forest ecosystems. This paper showed that the average FCO<sub>2</sub> of each type of forest in China was close to the average reported in similar forest ecosystems in other regions around the world. The results also showed that an increase in air temperature can greatly increase root respiration and microbial activity in these regions, leading to a high FCO<sub>2</sub>, thus contributing to an increase in the greenhouse effect. Dahikar and Rode (2014) used a MLP neural network trained by backpropagation to predict agricultural crop yield. In this work, they also considered several soil parameters, such as the type of soil, the pH of the soil, and the nitrogen content. Deneshkumar et al. (2015) presented a multilayer feedforward neural network for agricultural forecasting and compared the results with the traditional autoregressive integrated moving average (ARIMA) of Box and Jenkins, concluding that the ANN provides an attractive alternative. Pandey and Mishra (2017) used two different neural networks (GRNN and radial basis neural network [RBF]) for evaluating the yield of potatoes. According to their conclusion, both ANNs perform accurately, but the GRNN provides better results than the RBF. Many other studies can be found in the literature; however, the application proposed in this work is a novelty.

A soil's CO<sub>2</sub> emission is the result of complex interactions between the physical and biological transport phenomena of gas production in the soil, which are influenced by various physical and chemical properties of soils. The objective of this study is to investigate the use of a MLP neural network with BP (backpropagation algorithm) to predict patterns of spatiotemporal variability for short periods of soil CO<sub>2</sub> emission in areas of sugarcane using a green harvest system in the southeastern section of the State of São Paulo, considering the physical and chemical attributes of the soil.

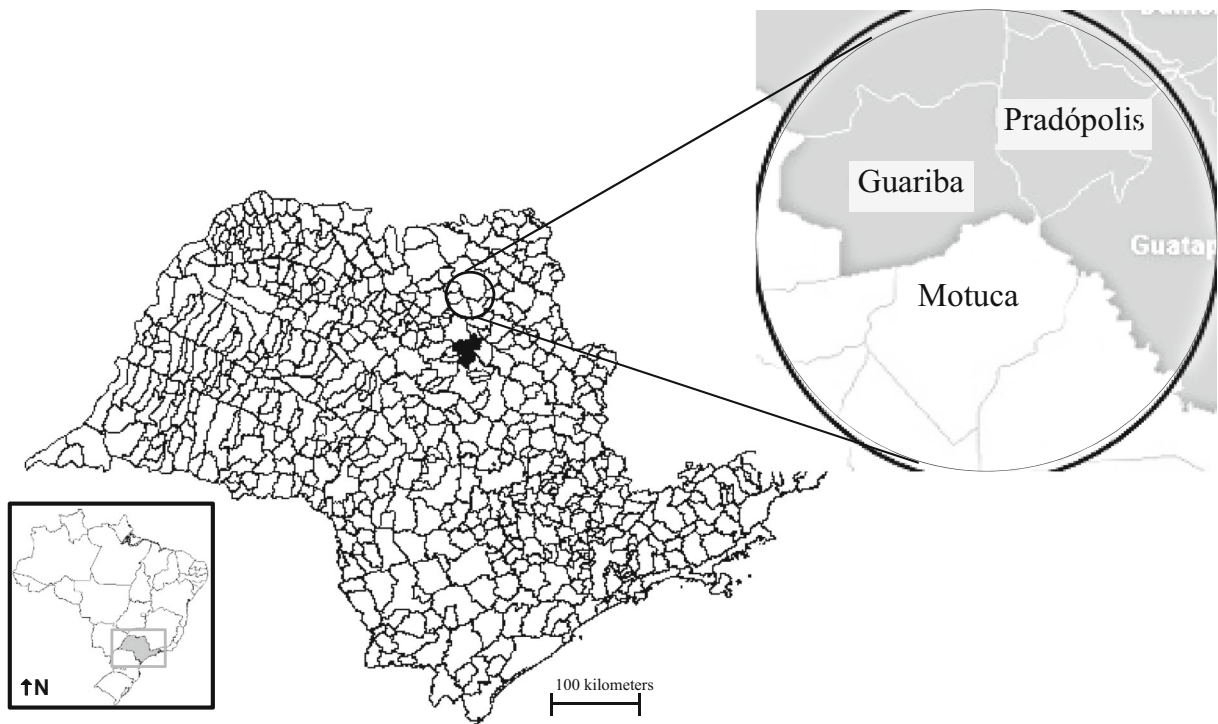
## Data collect and experience

The present study compiles the previous findings regarding the emission of CO<sub>2</sub> from soil from experiments conducted in three commercial areas in São Paulo State,

Brazil. The data originated from sugarcane areas in the green harvest, called green cane management, without burning practices, using mechanical harvesting that leaves large amounts of crop residues on the soil surface (mean of 12 tons ha<sup>-1</sup> year<sup>-1</sup>). These areas are as follows: Motuca (7 years of conversion from burned to green management), where the evaluation was carried out on October 25 and November 17 of 2008; Guariba (8 years of conversion from burned to green management), where the evaluation was carried out on July 14 and July 26 of 2010; and Pradópolis (15 years of conversion from burned to green management), where the evaluation was carried out on August 27 and September 14 of 2012 (Fig. 1). The soil CO<sub>2</sub> emission, soil temperature, and soil moisture were evaluated along short periods using sample grids, from which soil physical and chemical attributes were also evaluated (Panosso et al. 2012; Bicalho et al. 2014). The soils of all experimental areas were classified as Oxisol (Santos et al. 2013) and Eutruxox (USDA Soil Taxonomy). The climate, labeled Aw (according to Köppen), has a rainy tropical summer and a dry winter. The average rainfall is 1425 mm, concentrated between the months of October and March, with an average annual temperature of 22.2 °C recorded in the region for the past 30 years.

The CO<sub>2</sub> soil emission was registered by the LI-COR system (LI-8100), after the harvesting operation and during the early stages of crop growth. The soil temperature was monitored simultaneously with the CO<sub>2</sub> emission evaluation, using a temperature sensor that is part of the LI-8100 system. The soil moisture was determined with TDR equipment (time-domain reflectometry—Hydrosense TM, Campbell Scientific, Australia). Following the measurement period, soil samples were collected from 0 to 0.10 m in depth. The following routine tests were carried out: determining the content of organic matter (OM), the available P, K, Ca, Mg, and H + Al, and establishing the calculation of the sum of the bases and the exchange cation capacity (Van Raij et al. 2001). The bulk density was determined in undisturbed soil samples (EMBRAPA 1997). The total pore volume was distributed by size and macro- and micro-porosity and was determined using a fritted glass funnel under a pressure of a 60 cm height of water column (EMBRAPA 1997).

Data from 2012 (observed and estimated by the network) was used to analyze the experimental variogram (Webster and Oliver 1990), so that the spatial variability patterns of FCO<sub>2</sub> could be characterized, by



**Fig. 1** Map showing the site locations of the studies, Motuca in 2008, Guariba in 2010, and Pradópolis in 2012

estimating the semi-variance in a given distance. This estimate was determined by Eq. (1):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (1)$$

where  $h$  is the distance between pairs of points;  $N(h)$  is the number of pairs of points separated by the distance  $h$ ;  $Z(x_i)$  is the  $\text{FCO}_2$  value at the point  $x_i$ ; and  $Z(x_i + h)$  is the  $\text{FCO}_2$  value at the point  $x_i + h$ .

The choice of the model that best fits the experimental variograms (observed and estimated by the network) was made by cross-validation techniques, which involves the removal of each observation pertaining to the data set by estimating the value and interpolation method (data not shown). The chosen model best predicted the observed values, so that a linear regression equation between these values was produced, depending on the estimated values closer to the bisect—intercept equal to zero and angular coefficient equal to unity (Isaaks and Srivastava 1989). Theoretical models were used, as in Eqs. (2) to (4):

Exponential model:

$$\hat{\gamma}(h) = C_0 + C_1 \{1 - \exp[-(3h/a)]\} \quad (2)$$

where  $h > 0$ .

Spherical model:

$$\hat{\gamma}(h) = C_0 + C_1 \left[ \frac{3}{2}(h/a) - \frac{1}{2}(h/a)^3 \right] \quad (3)$$

where  $0 < h < d$ ,  $\hat{\gamma}(h) = C_0 + C_1$ , and  $h > a$ .

Gaussian model:

$$\hat{\gamma}(h) = C_0 + C_1 \{1 - \exp[-3(h/a^2)]\} \quad (4)$$

where  $0 < h < d$ , and  $d$  is the maximum distance at which the variogram was defined.

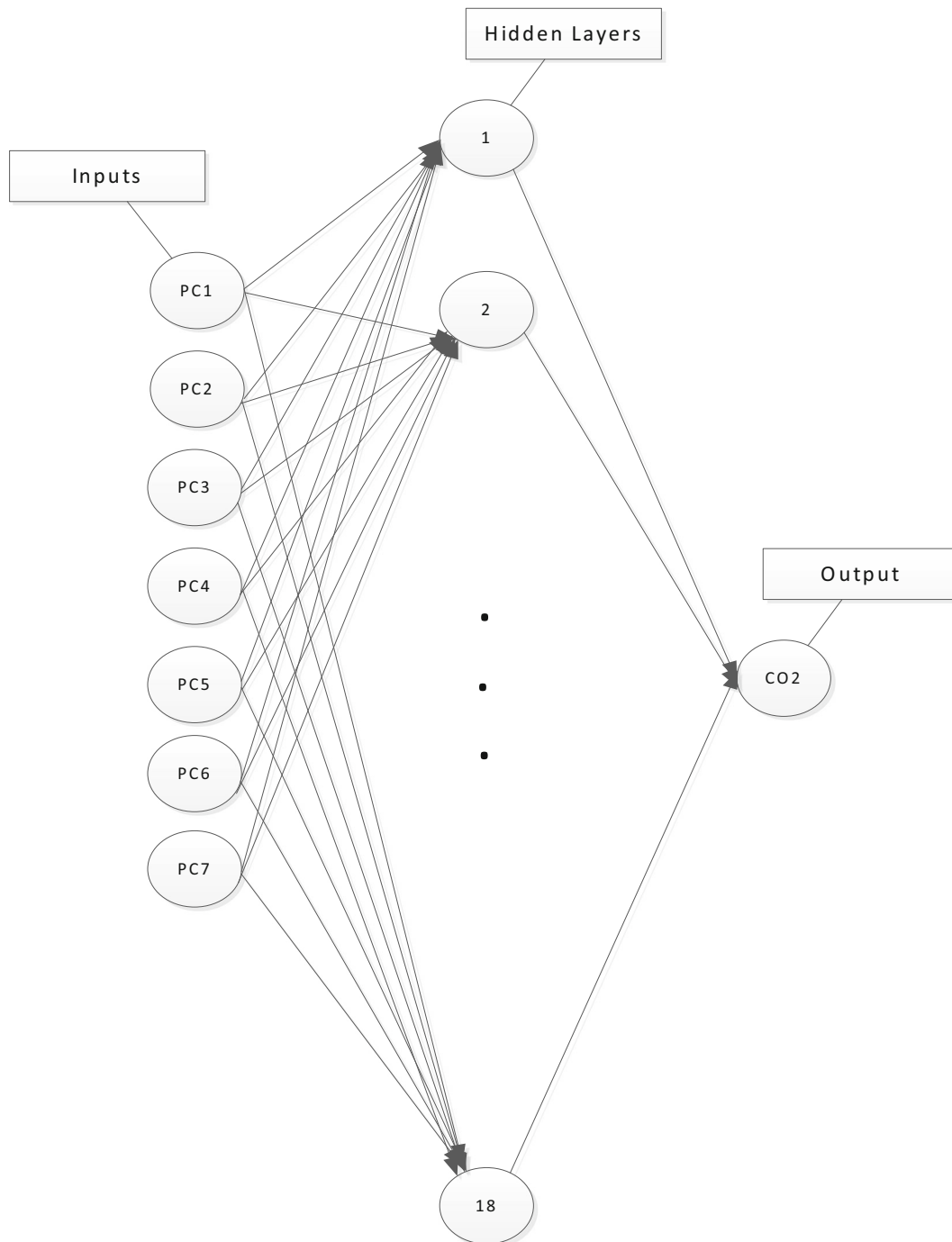
There are three important parameters in the variogram: nugget ( $C_0$ ); sill ( $C_0 + C_1$ ); and range ( $a$ ). The semi-variance value found at the intersection with the  $Y$ -axis, symbolized by  $C_0$ , is defined as the nugget and is the variability of  $\text{FCO}_2$  in smaller distances than those sampled. Ideally, the amount of semi-variance would be zero when the separation distance  $h$  is zero. In practice, this usually does not occur due to experimental errors and the occurrence of spatial variability in distances smaller than those used in the samples' grid (Wang et al. 2002). The sill ( $C_0 + C_1$ ) is the point at which the variogram model reaches a plateau and describes the variation of style. The range ( $a$ ) is the distance at which the landing occurs and from which there is spatial independence and no spatial autocorrelation



between observations. It is expected that differences in  $Z(x_i) - Z(x_i + h)$  progressively decrease so that  $h$  decreases, i.e.,  $\text{FCO}_2$  values (observed and estimated) located next to each other are more like each other than

those separated by great distances. Thus, it is expected that  $\hat{\gamma}(h)$  increases with the distance ( $h$ ).

The parameters of the models fitted to the experimental variograms that were used to estimate  $\text{FCO}_2$  in unsampled



**Fig. 2** Architecture of the MLP neural network

**Table 1** Real outputs obtained through the MLP neural network representing the identification of CO<sub>2</sub> emissions

Am	SRM	SO	Am	SRM	SO
500	0.126	0.087	510	0.103	0.085
501	0.110	0.071	511	0.195	0.167
502	0.126	0.082	512	0.113	0.108
503	0.100	0.080	513	0.111	0.066
504	0.072	0.072	514	0.402	0.369
505	0.143	0.082	515	0.194	0.099
506	0.171	0.166	516	0.091	0.107
507	0.220	0.160	517	0.135	0.097
508	0.130	0.086	518	0.083	0.078
509	0.169	0.092	519	0.103	0.131

locations were used in the construction of spatial pattern maps with the ordinary kriging technique, Eq. (5):

$$Z^*(x_0) = \sum_{i=1}^N \lambda_i Z(x_i) \quad (5)$$

where  $z^*$  is the value to be estimated at the point not sampled  $x_0$ ;  $N$  is the number of measured values;  $Z(x_i)$  is involved in the estimate; and  $\lambda_i$  are the weights associated with each measured value  $Z(x_i)$ .

This study used 2377 recorded observations in 2008, 2010, and 2012, and were divided into two groups:

- 1) 1314 registered data points from 2008 and 2010 were chosen randomly and used in the neural network training phase.
- 2) 1063 registered data points from 2012 were used for the diagnostic phase of the neural network, or

neural network classification, i.e., the estimation of FCO<sub>2</sub>.

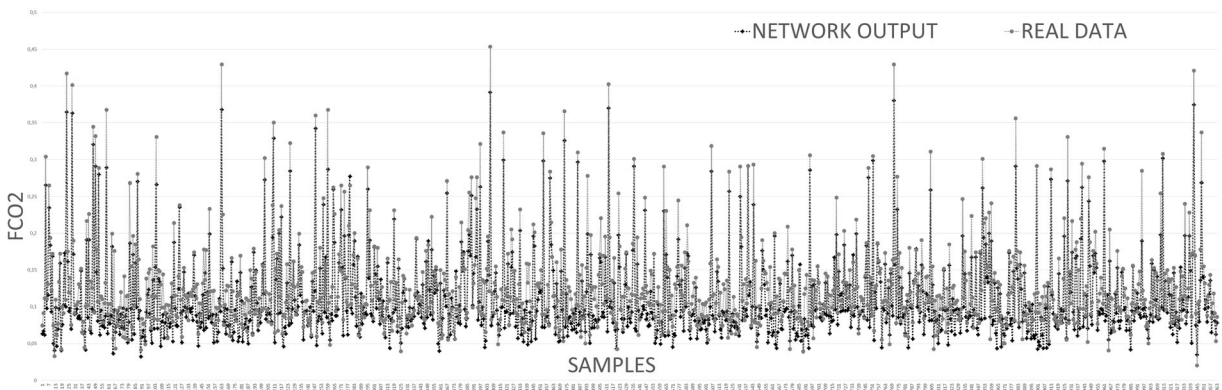
The neural network used had three layers. The input layer was composed of 7 neurons, the intermediate layer had 18 neurons, and the output layer had only 1 neuron corresponding to the neural network diagnosis, i.e., the soil CO<sub>2</sub> emission as shown in Fig. 2.

The neural network input variables were determined by a principal component analysis (PCA), whose main objective was to identify, in a data set, a variable that can explain a significant part of the variance observed in the system through linear correlations (Lawless and Heymann 2010). The PCA is a technique that transforms a large number of variables into a smaller set of non-correlated variables without a loss of information related to the initial variables, which are called the principal components (PCs). This smaller set of variables is generated by linear combinations of the original variables through the eigenvalues of the covariance matrix. These PCs work as axes, and each group applied to the component's equation yields the coordinates (scores), from which one can obtain its location in space. From a principal component, it is possible to obtain  $k$  variables, so:

$$\begin{aligned} PCp &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\ \text{Unity}(1) &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\ \text{Unity}(2) &= a_{11}X_1 + a_{21}X_2 + a_{p1}X_p \\ &\vdots \\ \text{Unity}(k) &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \end{aligned} \quad (6)$$

Considering the samples originating from the sugarcane area: organic matter (OM) content, cation exchange capacity (CEC), carbon stock (CS), average

### CO2 FLOW EMISSION REAL AND ESTIMATED



**Fig. 3** Results from the 1063 data samples of soil CO<sub>2</sub> emission (FCO<sub>2</sub>) presented by the neural network

temperature (avgT), average moisture (avgH), average free water pore space (avgPLA), hydrogenionic potential (pH), bulk density (BD), and total pore volume of the soil (TPV), were determined to be the principal components PC<sub>1</sub>, PC<sub>2</sub>, PC<sub>3</sub>, PC<sub>4</sub>, PC<sub>5</sub>, PC<sub>6</sub>, PC<sub>7</sub>, respectively, and were the inputs of the neural network. The input data were normalized between [0, 1] in order to avoid saturation of the neural network.

For an accurate analysis of CO<sub>2</sub> emission, the mean absolute percentage error (MAPE) and the coefficient of determination  $R^2$  were determined, respectively, by Eqs. (7) and (8):

$$\text{MAPE} = \frac{1}{N} \left\{ \sum_{i=1}^N (|ER - EO_{\text{MLP}}|) / ER \right\} \times 100 \quad (7)$$

$$R^2 = \sum_{i=1}^N (EO_{\text{MLP}} - ER_{\text{avg}})^2 / \sum_{i=1}^N (ER - ER_{\text{avg}})^2 \quad (8)$$

where  $ER$  = real CO<sub>2</sub> emission registered;  $EO_{\text{MLP}}$  = CO<sub>2</sub> emission from the MLP;  $ER_{\text{avg}}$  = real CO<sub>2</sub> emission average; and  $N$  is the number of samples.

## Results and discussion

This work used the structure of a neural network (MLP feedforward neural network) composed of three layers (input, hidden, output) with 7, 18, and 1 neurons, respectively. The basic training algorithm chosen is backpropagation algorithm, which algorithm is presented in the [Appendix](#). This simulation used the following parameters (see [Appendix](#)):

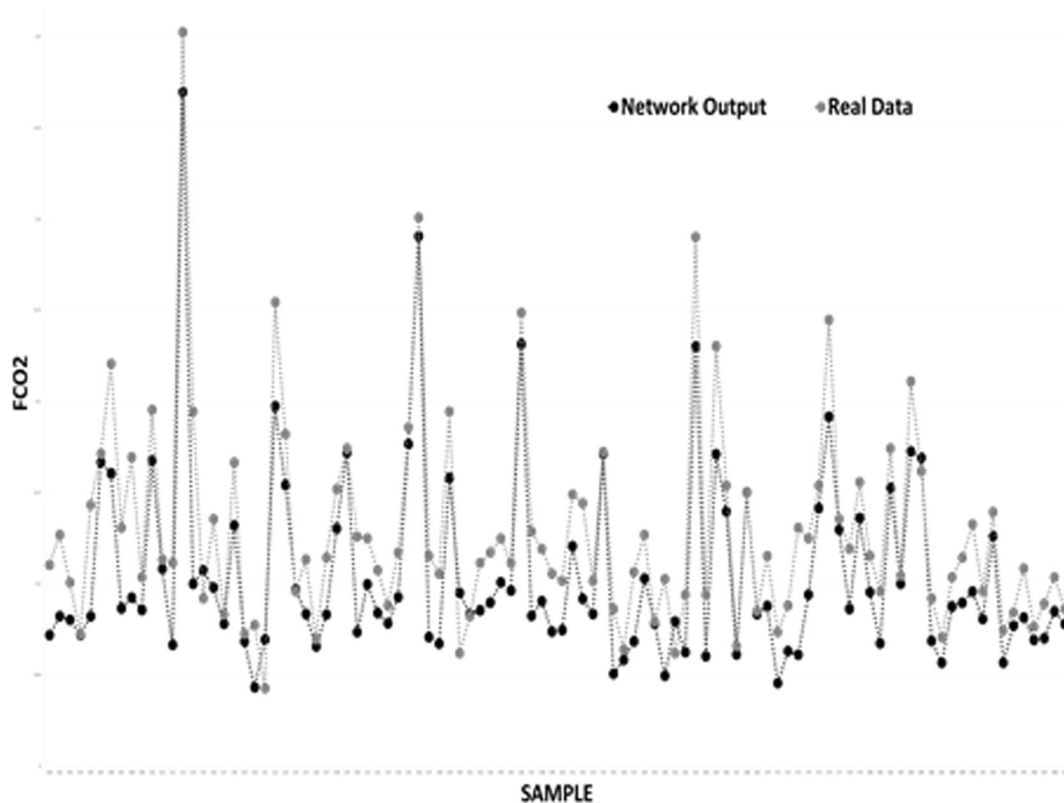
$\lambda = 0.3$  (inclination of sigmoid function)

$\gamma = 0.1$  (training rate)

and  $\eta = 0$  (momentum rate)

Another training algorithm, such as the Levenberg-Marquardt (Liu 2010), was also used to comparison, as well as another neural network called general regression neural network (GRNN) (Specht 1991).

The MLP neural network output was represented by the average soil CO<sub>2</sub> emission. Table 1 shows a small sample of the 1063 observations in the database for 2012 and the output obtained from the MLP neural network classification stage (diagnostic phase). The sample



**Fig. 4** Results from the 100 data samples of soil CO<sub>2</sub> emission (FCO<sub>2</sub>) presented by the neural network



(Am) was used to compare the measured real output data (SRM) with the obtained outputs (SO). The real outputs are the information obtained through the instruments during the experimental phase. The obtained outputs are the information estimated by the network proposed in this paper. Figure 3 shows the 1063 actual results from the application and the results obtained by the neural network for a sample of data collected from the rating group, and Fig. 4 presents 100 data samples (samples numbering from 500 to 599) to facilitate the data reading and the analysis of the behavior of the neural network. The precision analysis of the determined CO<sub>2</sub> emissions represents a MAPE of 18.2852% and a coefficient of determination  $R^2$  of 0.9188. When comparing to the Levenberg-Marquard training algorithm, MAPE is equal to 34.2984%, and in GRNN neural network, MAPE is equal to 43.0225%.

The Levenberg-Marquardt training is faster than backpropagation but depending on the size of the Hessian matrix, it has a disadvantage due to the need to invert this matrix, which demands time.

GRNN neural network, in general, is also very fast, but it was not adequate to solve the present problem when comparing the performance analyzing the MAPE.

That is the reason to choose backpropagation algorithm as the best one for the present problem (CO<sub>2</sub> emissions).

The emission of soil CO<sub>2</sub> varies in space and time, so the estimates of spatial dependence FCO<sub>2</sub> for the days studied in 2012 are presented in Table 2. The models fitted to the experimental variogram FCO<sub>2</sub> were mostly spherical, except for day 240, the first day of the evaluation, whose spatial variability structure was best described by the Gaussian model. The estimated values of soil CO<sub>2</sub> emission were similarly described in their spatial variability structure by spherical models, except for day 240 (exponential) and days 247 and 251 (Gaussian). The separation distances that comprise the autocorrelation data for the observed FCO<sub>2</sub> ranged from 8.34 m (day 255) to 21.66 m (day 249). Similar results were observed for the FCO<sub>2</sub> data estimated by the network, for which the parameter values varied from 6.34 m (day 255) to 23.2 m (day 249). Several models are used to describe the spatial variability of FCO<sub>2</sub>, with spherical models (La Scala et al. 2000; Kosugi et al. 2007; Konda et al. 2008; La Scala et al. 2010; Herbst et al. 2012; Tavares et al. 2016) or alternations

between spherical and exponential models (Tedeschi et al. 2006; Ohashi and Gyokusen 2007). The higher mean range value of the spatial variability structure indicates a more homogeneous distribution of FCO<sub>2</sub>. La Scala et al. (2000), valuating the FCO<sub>2</sub> in a grid of 100 × 100 m in a bare soil, obtained range values from 20.2 to 58.4 m. Teixeira et al. (2011) observed values ranging from 22.39 to 54.31 m with smaller sample sizes (60 × 60 m) in areas with soil surface residues. Changes in the range of spatiotemporal variability models of FCO<sub>2</sub> have been observed between seasons (Ohashi and Gyokusen 2007), between months (Stoyan et al. 2000), after precipitation events (La Scala et al. 2000), and even according to the grid sample size (Konda et al. 2008; Herbst et al. 2009).

Analyzing the data obtained from the observed FCO<sub>2</sub> and the estimated FCO<sub>2</sub>, it is possible to verify that the coefficient of the nugget effect has

**Table 2** Models and estimates of parameters adjusted to the experimental variograms of soil CO<sub>2</sub> emission (observed and estimated by the network), on different days of assessment 2012

Day July	Model	$C_0$	$C_0 + C_1$	$a$ (m)	$r^2$
FCO <sub>2</sub> observed					
240	Gau	0.119	1.197	14.96	0.72
242	Esf	0.066	0.148	11.6	0.75
244	Esf	0.058	0.159	19.98	0.85
247	Esf	0.063	0.150	19.16	0.79
249	Esf	0.113	0.321	21.66	0.86
251	Esf	0.111	0.239	17.27	0.72
255	Esf	0.050	0.128	8.34	0.68
258	Esf	0.052	0.140	17.12	0.80
Average	Esf	0.1179	0.272	18.87	0.67
FCO <sub>2</sub> estimated					
240	Exp	0.058	0.187	11.10	0.77
242	Esf	0.040	0.123	11.53	0.75
244	Esf	0.059	0.143	20.16	0.73
247	Gau	0.053	0.123	15.32	0.83
249	Esf	0.129	0.260	23.20	0.73
251	Gau	0.116	0.203	13.59	0.75
255	Esf	0.054	0.116	6.34	0.53
258	Esf	0.044	0.114	18.20	0.64
Average	Esf	0.047	0.211	19.32	0.77

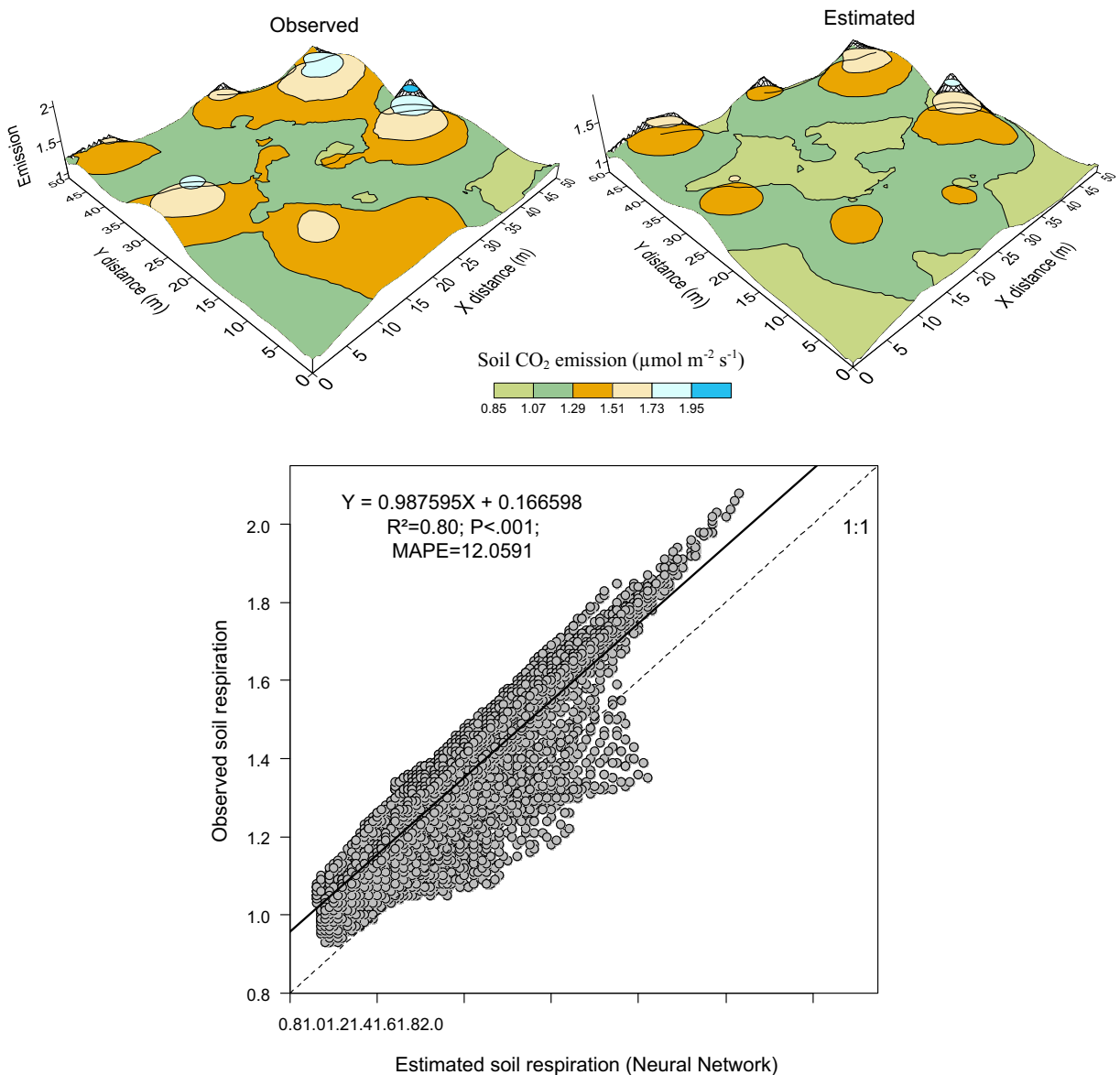
Exp exponential, Esf spherical, Gau Gaussian,  $r^2$  coefficient of determination of the model fit

moderate spatial dependence, i.e., the random component is important.

Figure 5 shows maps of the spatial pattern of the observed soil CO<sub>2</sub> flow and the values estimated by the neural network, and these data are of considerable similarity. By analyzing the grade, it is possible to see that the characteristics that point to higher emissions as well as to lower CO<sub>2</sub> emissions have been identified by the neural network with similar values.

The used network achieved satisfactory results in determining FCO<sub>2</sub> in sugarcane areas in a green harvest system, considering the diversity presented by the data and the fact that the results were determined by the main components according to the obtained MAPE and the correlation coefficient.

The neural network completed the learning process with data collected in 2008 and 2010 in different months and days and classified the data



**Fig. 5** Maps of the spatial pattern of soil CO<sub>2</sub> flux (FCO<sub>2</sub>) observed and estimated together with a map comparison scatterplot between observed and estimated FCO<sub>2</sub> values in kriging maps

collected in 2012. Even with the interference from several natural factors in this period, the network obtained a good prediction of the actual values, which shows that the capacity of the neural network for learning and adaptation is ideal for forecasting the CO<sub>2</sub> emission.

## Conclusions

Forecasting FCO<sub>2</sub> with artificial neural networks in sugarcane areas can contribute significantly to soil management adjustments to reduce the emission of CO<sub>2</sub> into the atmosphere and thus reduce the additional greenhouse effect. In addition, the use of this methodology can be a tool to reduce the uncertainties associated with estimates of this important component of the carbon cycle in the atmosphere.

Depending on the nature of the application and the consistency of the input data standards, it is possible for a network to train well and obtain excellent estimates. Thus, these neural networks can be applied to problems of dynamics and nonlinear relationships. They supply an analytical alternative to conventional techniques that are limited, often by strict assumptions of normality, linearity, and variable independence.

The FCO<sub>2</sub> forecasting results obtained by the MLP network and backpropagation algorithm are close to the experimentally measured values, proving an excellent efficiency of the use of artificial neural networks in agriculture.

**Author contributions** A.D.P.L., A.R.P., L.B.C., C.R.M., M.L.M.L., N.L.S.J., and R.L.B.F. contributed to the conception and/or design of the work. P.S.F., A.D.P.L., A.R.P., C.R.M., M.L.M.L., N.L.S.J., and R.L.B.F. participated in the conduction of the experiments and the acquisition of data. A.R.P. and A.D.P.L. performed data analysis. All authors contributed to the interpretation of data and to the drafting and revision of the manuscript, and approved the final version to be published.

**Funding information** The authors gratefully thank the funding and support from CAPES, CNPq, and FAPESP (08/58187-0; 10/20364-9; and 13/24926-0).

## Compliance with ethical standards

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

## Backpropagation algorithm

The initial weights are usually adopted as random numbers (Widrow and Lehr 1990). The backpropagation (BP) algorithm consists of adapting the weights such that the network quadratic error is minimized. The sum of the instantaneous quadratic error of each neuron of the last layer (network output) is given by (Widrow and Lehr 1990):

$$\varepsilon^2 = \sum_{i=1}^{no} \varepsilon_i^2 \quad (9)$$

where

$\varepsilon_i$   $d_i - y_i$ ;  
 $d_i$  desired output of the  $i$ th element of the last layer;  
 $y$  output of the  $i$ th element of the last layer; and  
 $no$  number of neurons of the last layer.

Considering the  $i$ th network neuron and using the descent gradient method, the weight adjustments can be formulated by (Widrow and Lehr 1990):

$$\Gamma_i(r+1) = \Gamma_i(r) + \theta_i(r) \quad (10)$$

where

$\theta_i(r)$   $-\gamma [\nabla_i(r)]$ ;  
 $\gamma$  stability control parameter or training rate;  
 $\nabla_i(r)$  quadratic error gradient related to neuron  $i$  weights; and  
 $\Gamma_i$  vector containing neuron  $i$  weights  
 $= [w_{0i}, w_{1i}, w_{2i}, \dots, w_{ni}]^T$ .

The adopted direction in Eq. (10) to minimize the objective function of the quadratic error corresponds to the gradient opposite direction. The  $\gamma$  parameter determines the vector length  $\phi_i(r)$ . The sigmoid function is defined by (Widrow and Lehr 1990):

$$y_i \Delta y_i(\lambda, \vartheta_i) = \left\{ \left( 1 + \exp(-\lambda \vartheta_i) \right) \right\} / \left\{ \left( 1 + \exp(-\lambda \vartheta_i) \right) \right\} \quad (11)$$

or

$$y_i \Delta y_i(\lambda, \vartheta_i) = 1 / \left\{ \left( 1 + \exp(-\lambda \vartheta_i) \right) \right\} \quad (12)$$

where

$\lambda$  constant that determines the slope of the function  $y_i$ .

The variation of Eqs. (11) and (12) are  $(-1, +1)$  and  $(0, +1)$ , respectively.

Next, calculating the gradient as shown in Eq. (1012) and considering the sigmoid function defined by Eq. (11) or (12) and the momentum term, the following adaptation weight scheme is obtained (Lopes et al. 2003):

$$\Pi_{ij}(r+1) = \Pi_{ij}(r) + \Delta\Pi_{ij}(r) \quad (13)$$

where

$$\Delta\Pi_{ij}(r) = 2\gamma(1-\eta)\beta_j x_i + \eta\Delta\Pi_{ij}(r-1); \quad (14)$$

$\Pi_{ij}$  weight corresponding to the connection with the  $i$ th and the  $j$ th neuron;

$\gamma$  training rate; and

$\eta$  momentum constant ( $0 \leq \eta < 1$ ).

If the  $j$ th element is in the last layer, then:

$$\beta_j = \sigma_j \varepsilon_j \quad (15)$$

where

$\sigma_j$   $\triangleq$  sigmoid function derivative, given by Eq. (11) or (12), respectively, related to  $v_j$

$$= \lambda/2(1-y_j^2). \quad (16)$$

$$= \lambda y_j(1-y_j^2). \quad (17)$$

If the  $j$ th element is in other layers, we have:

$$\beta_j = \sigma_j \sum_{k \in \Gamma(j)} w_{jk} \beta_k \quad (18)$$

$\Gamma(j)$  set of the element indices that are in the next layer to the  $j$ th element layer and are interconnected to the  $j$ th element.

The  $\gamma$  parameter that is used as a stability control for the iterative process is dependent on  $\lambda$ . The network weights are randomly initialized from the interval  $[0,1]$ . For convenience, the

parameter  $\gamma$  (training rate) can be redefined by the following (Lopes et al. 2003):

$$\gamma = \gamma^* / \lambda \quad (19)$$

Replacing Eq. (19) in Eq. (12) will “cancel” the amplitude dependency of  $\sigma$  related to  $\lambda$ . The  $\sigma$  amplitude will be maintained constant to every  $\lambda$ . This alternative is important considering that  $\lambda$  will only actuate in the left and right tails of  $\sigma$ . Equation (14) can then be written as the following:

$$\Delta\Pi_{ij}(r) = \{2\gamma^*(1-\eta)\beta_j/\lambda\}x_i + \eta\Delta\Pi_{ij}(r-1). \quad (20)$$

The BP algorithm executes as follows (Widrow and Lehr 1990):

- Step 1. Present a pattern  $X$  to the network, which provides an output  $Y$ .
- Step 2. Calculate the error (difference with the desired value and the output) for each output.
- Step 3. Determine the backpropagated error by the network associated with the partial derivative of the quadratic error.
- Step 4. Adjust the weights of each element.
- Step 5. Present a new pattern to the network and repeat the process until the convergence is attained (according to a predefined tolerance).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Bicalho, E. S., Panosso, A. R., Teixeira, D. D. B., Miranda, J. G. V., Pereira, G. T., & La Scala, N. (2014). Spatial variability structure of soil CO<sub>2</sub> emission and soil attributes in a sugarcane area. *Agriculture Ecosystems & Environment*, 189, 206–215. <https://doi.org/10.1016/j.agee.2014.03.043>.
- Cerri, C. E. P., Sparovek, G., Bernoux, M., Easterling, W. E., Melillo, J. M., & Cerri, C. C. (2007). Tropical agriculture and global warming: impacts and mitigation options. *Scientia Agricola*, 64, 83–99. <https://doi.org/10.1590/S0103-0162007000100013>.
- CONAB 2018. Available at: <https://www.conab.gov.br/info-agro/safras/graos>. Accessed date: 6 November 2018.
- Dahikar, S. S., & Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, 2(1), 683–686.
- Deneshkumar, V., Kannan, S. and Manikandan, M. (2015). *Designing of neural network models for agricultural*

- forecasting. Available online at <https://doi.org/10.1080/09720510.2015.1040237>, pp. 547–559.
- EMBRAPA. (1997). Brazilian Agricultural Research Corporation. In *Manual of soil analysis methods* (2nd. ed.). Brasília: MAPA (In Portuguese).
- Epron, D., Bosc, A., Bonal, D., & Freycon, V. (2006). Spatial variation of soil respiration across a topographic gradient in a tropical rain forest in French Guiana. *Journal of Tropical Ecology*, 22, 565–574. <https://doi.org/10.1017/S0266467406003415>.
- Figueiredo, E. B., & La Scala Júnior, N. (2011). Greenhouse gas balance due to the conversion of sugarcane areas from burned to green harvest in Brazil. *Agricultura, Ecosystems & Environment*, 141, 77–85. <https://doi.org/10.1016/j.agee.2011.02.014>.
- Haykin, S. (1999). *Neural networks: a comprehensive foundation*. New Jersey: Prentice-Hall.
- Herbst, M., Prolingheuer, N., Graf, A., Huisman, J. A., Weihermuller, L., & Vanderborght, J. (2009). Characterization and understanding of bare soil respiration spatial variability at plot scale. *Vadose Zone Journal*, 8, 762–771.
- Herbst, M., Bornemann, L., Graf, A., Welp, G., Vereecken, H., & Amelung, W. (2012). A geostatistical approach to the field-scale pattern of heterotrophic soil CO<sub>2</sub> emission using co-variables. *Biogeochemistry*, 111, 377–392.
- IPCC - Intergovernmental Panel on Climate Change. (2007). *Climate change 2007: the physical science basis. Summary for policymakers*, Geneva, Switzerland. Available in: [http://www.ipcc.ch/publications\\_and\\_data/publications\\_ipcc\\_fourth\\_assessment\\_report\\_wg1\\_report\\_the\\_physical\\_science\\_basis.htm](http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm).
- IPCC - Intergovernmental Panel on Climate Change. (2014). *Climate change 2014: synthesis report for policymakers*, Geneva, Switzerland. Available in: [http://www.ipcc.ch/pdf/assessment-report/ar5/wg3/ipcc\\_wg3\\_ar5\\_full.pdf](http://www.ipcc.ch/pdf/assessment-report/ar5/wg3/ipcc_wg3_ar5_full.pdf). Access in 23 abr. 2016.
- Isaaks, E. H., & Srivastava, R. M. (1989). *Applied geostatistics*. Nova York: Oxford University Press.
- Kartalopoulos, S. V. (1996). *Understanding neural networks and fuzzy logic: basic concepts and applications*. Piscataway, NJ, USA: IEEE Press.
- Kim, M., & Gilley, J. E. (2008). Artificial neural network estimation of soil erosion and nutrient concentrations in runoff from land application areas. *Computer and Electronics in Agriculture*, 64, 268–275.
- Konda, R., Ohta, S., Ishizuka, S., Arai, S., Ansori, S., Tanaka, N., & Hardjono, A. (2008). Spatial structures of N<sub>2</sub>O, CO<sub>2</sub>, and CH<sub>4</sub> fluxes from Acacia mangium plantation soils during a relatively dry season in Indonesia. *Soil Biology & Biochemistry*, 40, 3021–3030.
- Kosugi, Y., Mitani, T., Ltoh, M., Noguchi, S., Tani, M., Matsuo, N., Takanashi, S., Ohkubo, S., & Nik, A. R. (2007). Spatial and temporal variation in soil respiration in a Southeast Asian tropical rainforest. *Agricultural and Forest Meteorology*, 147, 35–47.
- La Scala, N., Marques, J., Pereira, G. T., & Cora, J. E. (2000). Short-term temporal changes in the spatial variability model of CO<sub>2</sub> emissions from a Brazilian bare soil. *Soil Biology & Biochemistry*, 32, 1459–1462.
- La Scala, N., de Sá Mendonça, E., Vanir de Souza, J., Panosso, A. R., Simas, F. N. B., & Schaefer, C. E. G. R. (2010). Spatial and temporal variability in soil CO<sub>2</sub>-C emissions and relation to soil temperature at King George Island, maritime Antarctica. *Polar Science*, 4, 479–487.
- Lawless, H. T., & Heymann, H. (2010). *Data relationships and multivariate applications, sensory evaluation of food—principles and practices* (pp. 433–449). Berlin: Springer.
- Lentzsch, P., Wieland, R., & Wirth, S. (2005). Application of multiple regression and neural network approaches for landscape-scale assessment of soil microbial biomass. *Soil Biology & Biochemistry*, 37, 1577–1580. <https://doi.org/10.1016/j.soilbio.2005.01.017>.
- Liu, H. (2010). On the Levenberg-Marquardt training method for feedforward neural networks. In *Proceedings of the 2010 International Conference on Natural Computation, Incn'10*, volume 1.
- Lopes, M. L. M., Minussi, C. R., & Lotufo, A. D. P. (2003). Electrical load forecasting formulation by a fast neural network. *Engineering Intelligent Systems*, 11(1), 51–57. <http://hdl.handle.net/11449/9730>.
- Lotufo, A. D. P., Lopes, M. L. M., & Minussi, C. R. (2007). Sensitivity analysis by neural networks applied to power systems transient stability. *Electric Power Systems Research*, 77, 730–738. <https://doi.org/10.1016/j.epsr.2005.09.020>.
- Luca, E. F., Feller, C., Cerri, C. C., Barthès, B., Chaplot, V., Campos, D. C., & Manechini, C. (2008). Evaluation of physical attributes and carbon and nitrogen stocks in soils with no burning of sugar cane. *Brazilian Journal of Soil Science*, 32, 789–800. (In Portuguese). <https://doi.org/10.1590/S0100-06832008000200033>.
- Merdun, H. (2011). Self-organizing map artificial neural network application in multidimensional soil data analysis. *Neural Computing and Applications*, 20, 1295–1303. <https://doi.org/10.1007/s00521-010-0425-1>.
- Moitinho, M. R., Padovan, M. P., Panosso, A. R., & La Scala, N. (2013). Effect of soil preparation and residue of the sugarcane harvest on the emission of CO<sub>2</sub>. *Brazilian Journal of Soil Science*, 37, 1720–1728. (In Portuguese). <https://doi.org/10.1590/S0100-06832013000600028>.
- Moretti, J. F., Minussi, C. R., Melges, J. L. P., Akasakil, J. L., & Tashima, M. M. (2016). Prediction of modulus of elasticity and compressive strength of concrete specimens by means of artificial neural networks. *Acta Scientiarum Technology*, 38, 65–70. <https://doi.org/10.4025/actascitechnol.v38i1.27194>.
- Ohashi, M., & Gyokusen, K. (2007). Temporal change in spatial variability of soil respiration on a slope of Japanese cedar (*Cryptomeria japonica* D. Don) forest. *Soil Biology & Biochemistry*, 39, 1130–1138.
- Pandey, A., & Mishra, A. (2017). Application of artificial neural networks in yield prediction of potato crop. *Russian Agricultural Sciences*, 43(3), 266–272.
- Panosso, A. R., Peillo, L. I., Ferraudo, A. S., Pereira, G. T., Miranda, J. G. V., & La, S. J. (2012). Fractal dimension and anisotropy of soil CO<sub>2</sub> emission in a mechanically harvested sugarcane production area. *Soil & Tillage Research*, 124, 8–16. <https://doi.org/10.1016/j.still.2012.04.005>.
- Rayment, M. B., & Jarvis, P. G. (2000). Temporal and spatial variation of soil CO<sub>2</sub> efflux in a Canadian boreal forest. *Soil*



- Biology and Biochemistry*, 32, 35–45. [https://doi.org/10.1016/S0038-0717\(99\)00110-8](https://doi.org/10.1016/S0038-0717(99)00110-8).
- Razafimbelo, T., Barthès, B., Larré-Larrouy, M. C., de Luca, E. F., Laurent, J. Y., Cerri, C. C., & Feller, C. (2006). Effect of sugarcane residue management (mulching versus burning) on organic matter in a clayey Oxisol from southern Brazil. *Agriculture Ecosystem & Environment*, 115, 285–289. <https://doi.org/10.1016/j.agee.2005.12.014>.
- Santos, H. G., Jacomine, P. K. T., Anjos, L. H. C., Oliveira, V. A., Oliveira, J. B., Coelho, M. R., Lumberras, J. F., & Cunha, T. J. F. (2013). *Brazilian system of soil classification*. Rio de Janeiro: Brazilian Agricultural Research Corporation (EMBRAPA) Soil.
- Song, X., Peng, C., Zhao, Z. S., Zhang, Z., Guo, B., Wang, W., Jiang, H., & Zhu, Q. (2014). Quantification of soil respiration in forest ecosystems across China. *Atmospheric Environment*, 94, 546–551. <https://doi.org/10.1016/j.atmosenv.2014.05.071>.
- Specht, D. F. (1991). A generalized regression neural network. *IEEE Transactions on Neural Networks*, 2, 568–576.
- Stoyan, H., De-Polli, H., Bohm, S., Robertson, G. P., & Paul, E. A. (2000). Spatial heterogeneity of soil respiration and related properties at the plant scale. *Plant and Soil*, 222, 203–214.
- Tavares, R. L. M., Siqueira, D. S., Panosso, A. R., Castioni, G. A. F., Souza, Z. M., & La Scala Jr., N. (2016). Soil management of sugarcane fields affecting CO<sub>2</sub> fluxes. *Scientia Agricola*, 7, 543–551.
- Tedeschi, V., Rey, A., Manca, G., Valentini, R., Jarvis, P. G., & Borghetti, M. (2006). Soil respiration in a Mediterranean oak forest at different developmental stages after coppicing. *Global Change Biology*, 12, 110–121.
- Teixeira, D. B., Panosso, A. R., Cerri, C. E. P., Pereira, G. T., & La Scala, N. (2011). Soil CO<sub>2</sub> emission estimated by different interpolation techniques. *Plant and Soil*, 345, 187–194.
- Teixeira, D. D. B., Bicalho, E. S., Panosso, A. R., Perillo, L. I., Iamaguti, J. L., Pereira, G. T., & La Scala, N. (2012). Uncertainties in the prediction of spatial variability of soil CO<sub>2</sub> emissions and related properties. *Brazilian Journal of Soil Science*, 36, 1466–1475. <https://doi.org/10.1590/S0100-6832012000500010>.
- Ussiri, A. N., & Lal, R. (2009). Long-term tillage effects on soil carbon storage and carbon dioxide emissions in continuous corn cropping system from an Alfisol in Ohio. *Soil & Tillage Research*, 104, 39–47. <https://doi.org/10.1016/j.still.2008.11.008>.
- Raij, B. Van, Andrade, J. C., Cantarella, H., and Quaggio, J. A. (2001). *Chemical analysis for the evaluation of tropical soils*. Agronomic Institute of Campinas. (In Portuguese).
- Wang, G., Gertner, G., Singh, V., Shinkareva, S., Parysow, P., & Anderson, A. (2002). Spatial and temporal prediction and uncertainty of soil loss using the revised universal soil loss equation: a case study of the rainfall-runoff erosivity R factor. *Ecological Modelling*, 153, 143–155. [https://doi.org/10.1016/S0304-3800\(01\)00507-5](https://doi.org/10.1016/S0304-3800(01)00507-5).
- Wasserman, P. D. (1989). *Neural computing—theory and practice*. New York: Van Nostrand Reinhold.
- Webster, R., & Oliver, M. A. (1990). *Statistical methods in soil and land resource survey*. New York: Oxford University Press.
- Werbos, P. J. (1974). *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. [PhD thesis]. [Harvard]: Harvard University.
- Widrow, B., & Lehr, M. A. (1990). 30 years of adaptive neural networks: perceptron, Madaline, and backpropagation. *Proceedings of the IEEE*, 78, 1415–1442. <https://doi.org/10.1109/5.58323>.