Environ Monit Assess (2023) 195:1074
https://doi.org/10.1007/s10661-023-11679-8

RESEARCH

Artificial neural networks and adaptive neuro-fuzzy inference systems for prediction of soil respiration in forested areas southern Brazil

Maria Elisa Vicentini · Paulo Alexandre da Silva · Kleve Freddy Ferreira Canteral · Wanderson Benerval De Lucena · Mario Luiz Teixeira de Moraes · Rafael Montanari · Marcelo Carvalho Minhoto Teixeira Filho · Nelson José Peruzzi · Newton La Scala Jr · Glauco De Souza Rolim · Alan Rodrigo Panosso

Received: 24 February 2023 / Accepted: 4 August 2023

Abstract The purpose of this study was to estimate the temporal variability of CO_2 emission (FCO_2) from O_2 influx into the soil (FO_2) in a reforested area with native vegetation in the Brazilian Cerrado, as well as to understand the dynamics of soil respiration in this ecosystem. The database is composed of soil respiration data, agroclimatic variables, improved vegetation index (EVI), and soil attributes used to train machine learning algorithms: artificial neural network (ANN) and an adaptive neuro-fuzzy inference system (ANFIS). The predictive performance was evaluated based on the mean absolute error (MEA), root mean square error (RMSE), mean absolute percentage error (MAPE), agreement index (d), confidence coefficient (c), and coefficient of determination (R^2). The best estimation results for validation were FCO_2 with multilayer perceptron neural network (MLP) ($R^2 = 0.53$, RMSE = 0.967 μmol m^{-2} s^{-1}) and radial basis function neural network (RBF) ($R^2 = 0.54$, RMSE = 0.884 μmol m^{-2} s^{-1}) and FO_2 with MLP ($R^2 = 0.45$, RMSE = 0.093 mg m^{-2} s^{-1}) and RBF ($R^2 = 0.74$, 0.079 mg m^{-2} s^{-1}). Soil temperature and macroporosity are important predictors of FCO_2 and FO_2. The best combination of variables for training the ANFIS was selected based on trial and error. The results were as follows: FCO_2 ($R^2 = 16$) and FO_2 ($R^2 = 29$). In all models, FCO_2 outperformed FO_2. A primary factor analysis was performed, and FCO_2 and FO_2 correlated best with the weather and soil attributes, respectively.

Keywords Soil CO_2 emission · Oxygen influx · Reforestation, Tropical ecosystems · Artificial intelligence

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10661-023-11679-8.

M. E. Vicentini (✉) · P. A. da Silva · K. F. F. Canteral · W. B. De Lucena · N. J. Peruzzi · N. La Scala Jr · G. De Souza Rolim · A. R. Panosso
Department Engineering and Exact Sciences, School of Agricultural and Veterinarian Sciences, São Paulo State University (FCAV/UNESP), Via de Acesso Prof. Paulo Donato Castellane S/N, Jaboticabal, São Paulo 14884-900, Brazil
e-mail: mevicentini@gmail.com

M. L. T. de Moraes
Department of Phytotecnics, Faculty of Engineer (FEIS/UNESP), Avenida Brasil–Centro, Ilha Solteira, São Paulo 15385-000, Brazil

R. Montanari · M. C. M. T. Filho
Department of Phytosanity, Rural Engineering and Soils, Faculty of Engineer (FEIS/UNESP), Avenida Brasil–Centro, Ilha Solteira, São Paulo 15385-000, Brazil

Introduction

The tropics contain the highest biodiversity on Earth, yet they are the hardest hit by anthropogenic actions (Barlow et al., 2018). Brazil has contributed to the loss of tropical biodiversity and currently taking the lead among the countries that have deforested the

majority of their primary forests. In recent decades, the removal of native vegetation in the Cerrado biome has increased considerably, reaching higher rates than those of deforestation in the Amazon region (Trigueiro et al., 2020). Highlighting the last 4 years (2019–2022), due to poor political management, Brazil has faced attacks on its environmental legislation, corroborating with the dismantling of its environmental policy, which has resulted in the advancement of deforestation and authorization for the destruction of the Areas of Urban Permanent Preservation (APP).

The Cerrado occupies approximately 24% of the Brazilian territory, and although it is a key component of the regional carbon (C) cycle, it is one of the most threatened biomes globally owing to the conversion of forests into agricultural areas (Strassburg et al., 2017). The biome stands out for its great biodiversity and heterogeneity of landscapes, which provide important ecosystem services; from a hydrological point of view, they are related to the large number of river sources that contribute to high water yield and the soil performs several functions such as filtering the water, nutrient cycling, and food production (Prado et al., 2016).

Notably, land use change (LUC) in this biome has also caused changes in the microbial community, which plays a fundamental role in the biogeochemical cycles of C, nitrogen (N), and phosphorus (P), as well as in maintaining soil quality (de Souza & Procópio, 2021). The degradation of vegetation cover and the replacement of native vegetation directly influences C and Nstocks, primarily by reducing the input of organic material into the soil, thereby decreasing soil quality (Nourbakhsh, 2007; Yuan et al., 2019).

Soil respiration (R_s) is the sum of autotrophic (root) and heterotrophic (microbes, soil fauna) respiration (Hanson et al., 2000). Basal (R_B) and substrate-induced respiration ($S_{RB)}$ is a major bioindicator of soil quality (Sparling, 1997). As carbon dioxide is produced and emitted during soil respiration (FCO_2), it is a major contributor to atmospheric CO_2. Furthermore, in this dynamic, oxygen is consumed (FO_2) as a result of the metabolic processes of microbial activity and root.

Especially, soil respiration varies with time and space and is influenced by soil water content and temperature (Adachi et al., 2006; Kim et al., 2012). In recent years, a number of studies have correlated R_s with meteorological data (Ouyang & Zheng, 2000; La Scala et al., 2003); oxygen availability (de Almeida et al., 2018), soil physical attributes such as class and pore size (de Oliveira Silva et al., 2019), the quantity, quality, and distribution of plant residues (de Araújo Santos et al., 2019), as well as with vegetation indices such as the enhanced vegetation index (EVI) (Huang et al., 2012). Specifically, Raich and Tufekcioglu (2000) observed that vegetation type, forest microclimate, and net primary productivity are positively correlated with R_s.

This is because vegetation and canopy age affect several soil properties such as the C/N ratio and pH (Lucas-Borja et al., 2019). Furthermore, the type of canopy cover (monoculture or mixed) has an effect on the physical, chemical, and biological characteristics of soil, as well as on the contribution of organic material. In addition, studies have shown that composite forests with different species increase microbial respiration (Kooch & Ghaderi, 2021; Kooch & Noghre, 2020) and influence soil CO_2 flux (Laganière et al., 2012).

The dynamics of O_2 influx are still poorly understood due to the limited number of studies, and there are some uncertainties, especially in tropical soils (de Almeida et al., 2018; Vicentini et al., 2019). The dynamics of gases, mainly oxygen and carbon dioxide, between the soil and atmosphere occur under the influence of both pressure and concentration gradients (Gliński & Stepniewski, 1985). Recently, De Lucena et al. (2023) observed that in tropical soils, the influx of O_2 is influenced by atmospheric pressure and soil temperature. At this moment, we are sure that research needs to be advanced in this area. In this context, it is important to estimate and understand the dynamics of FCO_2 and FO_2 using environmental variables and soil attributes. This especially applies to areas of reforestation with mixed vegetation, including the Cerrado.

The modeling via machine learning (ML) has increased rapidly, while the field of soil science has helped improve our understanding of the causes of soil variation temporal, the environmental variables that control soil distribution (Padarian et al., 2019), and the estimation of soil respiration (Bond-Lamberty, 2018; Ebrahimi et al., 2019; Freitas et al., 2018). Machine learning algorithms include artificial neural network (ANN) and adaptive network-based fuzzy inference system (ANFIS) models. Although ANFIS models have been successfully used to explain of the variability of R_s in different ecosystems (Zhao

et al., 2017), few studies have used ANFIS in the prediction of R_s in planted forest areas (Canteral, 2020; Canteral et al., 2023).

Concerning oxygen modeling, although the majority of estimates have focused on predicting dissolved oxygen concentration in water with ANN, the ANFIS system has shown good results (Heddam, 2014; Zounemat-Kermani et al., 2014; Lamzouri et al., 2017). However, some estimates have recorded large mean absolute percentage errors (MAPE) associated with these models (Liu et al., 2019). Notably, studies on soil O_2 influx in soil using machine learning are scarce.

Given the above, we hypothesize that of this study is that agroclimatic variables such as EVI and soil attributes can be used as predictor variables to model the temporal variability of CO_2 emissions and O_2 influx in reforested areas using machine learning techniques. The objective of the study is to evaluate and compare two ANN architectures, namely multilayer perceptrons (MLP) and radial basis function (RBF) networks and an ANFIS mode in predicting CO_2 emissions and O_2 influx in Cerrado areas reforested with native species.

Materials and methods

In the present study, a database was created based on the evaluation of the temporal variability of soil respiration (FCO_2 and FO_2) in an area that has suffered intense land use and occupation and is currently inhabited by native species. Area history: until 1977, the site is consisted of native vegetation typical of the Cerrado; and after the change in land use, the area was cultivated with rice (*Oryza sativa*). In 1986, with the construction of hydropower plants in the region, there is a need for reforestation at the site. Therefore, the study area was reforested with mixed native species. The readings of the FCO_2 and FO_2 fluxes were recorded between 2015 and 2016.

The mean FCO_2 was 4.059 μmol m^{-2} s^{-1}, the mean FO_2 was 0.371 mg m^{-2} s^{-1}, the soil temperature was 25.99 °C, and the soil water content (SWC) was 11.56% (Vicentini et al., 2019) (Table S1 in the Supplementary Material). Below is a description of how environmental variables were collected to form a robust database with 24 variables for the estimation of FCO_2 and FO_2 fluxes.

Study area

The study area was a riparian forest in the Cerrado, composed of the following 21 native species: *Albizia lebbeck*, *Chorisia speciosa*, *Cydonia oblonga*, *Enterolobium contortisiliquum*, *Eugenia jambolana*, *Ficus clusiifolia*, *Holocalyx glaziovii*, *Hovenia dulcis*, *Jacaranda semisenata*, *Koelreuteria paniculata*, *Moquilea tomentosa*, *Morus nigra*, *Myroxylon balsamum*, *Peltophorum dubium*, *Psidium guajava*, *Pterocarpus quercinus*, *Peltophorum vogelianum, Spondias venulosa, Tabebuia chrysotricha, Tabebuia impetiginosa,* and *Tabebuia odontodiscus*. The species were randomly distributed in a spacing of 3 × 2 m. The forest is located in the Fazenda Experimental da Faculdade de Engenharia (UNESP), in the city of Selvíria, Mato Grosso do Sul, Central Brazil, on the banks of the Paraná River (Fig. S1 in the Supplementary Material), with central coordinates of 20°20′53.41″ S and 51°23′55.50″ W, and an altitude of 354 m above sea level.

The climate of the region was classified as C1dAa′ by the Thornthwaite system (Rolim et al., 2007), with a mean annual temperature of 23.5 °C, mean annual rainfall of 1300 mm, and relative humidity between 70 and 80%. The local soil was classified as Dystrophic Red Latosol (LVd) according to the Brazilian Soil Classification System (dos Santos et al., 2022). The relief was moderately flat and undulating with 542.38 g kg^{-1} sand, 65.21 g kg^{-1} silt, and 392.39 g kg^{-1} clay in the 0–0.10 m layer. For approximately 10 years, the area was managed with rice cultivation.

Determination of soil CO_2 emission, O_2 influx, soil temperature, and water content

Soil CO_2 emissions were recorded between 07:00 and 12:00 using a LI-8100 portable system (LI-COR, NE, USA). Twenty-five sampling points were demarcated with polyvinyl chloride (PVC) collars (0.10 m diameter and 0.85 m height) and installed in the soil at a depth of 5 cm in each plot. The reading time was 120 s. Simultaneously, with the FCO_2 readings, soil temperature (T_s) and soil water content (SWC) were determined. Soil temperature was recorded for all studied points in the 0–0.12 m layer using a digital thermometer, while water content was evaluated using a portable TDR-Campbell® system (Hydrosense™, Campbell Scientific, Australia), consisting of a probe with two 20 cm rods, which were inserted into the

soil perpendicular to the surface. Further information was detailed by (Panosso et al., 2012).

The O_2 influx was determined with the aid of a portable UV Flux (0–25%) sensor (CO_2Meter, Inc., Ormond Beach, FL, USA) at ten selected points in the area. To determine the O_2 uptake in the soil, 10 points in each area were selected using the portable UV flux (0–25%) sensor (CO_2Meter, Inc). The O_2 influx into the soil was calculated by considering the atmospheric pressure, air temperature, and chamber volume. Further information was detailed by (de Almeida et al., 2018; Vicentini et al., 2019).

Soil data and environmental variables

Soil physical analyses were performed in the 0–0.10 m layer after the FCO_2 and FO_2 readings. The following results were observed: macroporosity = 0.049 (m^3 m^{-3}), microporosity = 0.394 (m^3 m^{-3}), and soil density = 1.34 (g dm^{-3}) (according to the methodology proposed by Embrapa (1997). For the chemical analyses, the following results were observed: phosphorus content (resin) = 6.16 (mg dm^{-3}), aluminum = 5.88 (mmolc dm^{-3}), soil organic matter = 25.2 (g dm^{-3}), pH (01 mol L^{-1} $CaCl_2$) = 4.44, cation exchange capacity (CEC) = 82.98, and base sum = 33.06 (mmolc dm^{-3}), according to the methodology proposed by (Raij et al., 2001). Litter samples were collected in May of 2016. The organic carbon and nitrogen contents were determined using the methods described by Tedesco et al. (1995) and Bataglia et al. (1983), respectively.

The meteorological database was obtained from the historical records of UNESP-CANAL CLIMA for the interval from November 2015 to October 2017. (http://clima.feis.unesp.br/index.php) including PP (mm) = rainfall, w (m/s) = wind speed, ET_o PN-M (mm/day) = reference evapotranspiration by Penman–Monteith, P_{atm} (kPa) = atmospheric pressure, T_{air} = air temperature, PAR (µmol m^{-2}) = photosynthetically active radiation, and RSG (MJ m^{-2} day^{-1}) = global solar radiation. Biophysical data EVI = spectral index; MODIS-EVI data were obtained using SATVeg—Temporary Vegetation Analysis System, developed by Embrapa Informática Agropecuária. The MODIS images used for the development of SATVeg were obtained from the LP DAAC Data Pool (http://lpdaac.usgs.gov). These images were composed of the maximum values of the vegetation indices every 16 days. The spatial resolution of the MODIS sensor for the product used was 250 m. To minimize the effects of time series noise, the data were obtained with pre-filtering and filtering available in the tool. The Savitzky-Golay filter (Savitzky & Golay, 1964) was selected. Previous studies have demonstrated the effectiveness of this filter in reducing the effect of noise in the MODIS time series, caused primarily by cloud contamination and atmospheric variability (Chen et al., 2004; Liu et al., 2022; Yang et al., 2021).

Data mining

In this paper, we develop regression models of FCO_2 and FO_2 from two feed-forward neural network architectures (multilayer perceptron and radial basis function) and the hybrid form of ANNs, i.e., adaptive neuro-fuzzy inference system. The dependent variables for modeling constituted $n = 500$ and $n = 175$ observations for FCO_2 and FO_2, respectively. In all ANN models, 70% of the data were used for calibration and 30% for validation.

Artificial neural network (ANN)

The architecture of a multilayer perceptron neural (MLP) consists of four main layers: input layer, two hidden layers, and output one intermediate layers are known as hidden layers between the input and output layers, which are trained by the backpropagation (BP) algorithm. The basic idea is to propagate the error signal calculated in the training step back to all neurons and the activation function of the sigmoidal type (Haykin, 2009).

The radial basis function neural (RBF) is alternatives to MLP, consisting of three layers: an input layer, a hidden layer (kernel), and an output layer. The number of neurons in the hidden layer can be adjusted according to actual needs. Learning can be supervised, unsupervised, or blended, but when training is supervised, the weights between the hidden layer and output layer are calculated to minimize the error between the desired and obtained output. The most common activation function of the RBF is Gaussian. For each estimate (i.e., FCO_2 and FO_2) of different

architectures, we observe that 100 epochs are sufficient for error convergence.

Step 1: We trained 100 radial basis function neural networks (RBF) and 100 multilayer perceptron neural networks (MLPs). Regarding the original data, the training set consisted of 24 variables, namely T_s, SWC, P, pH, H + Al, SB, CEC, macro, micro, Bd, Cstock, Nstock, CN, sand, silt, clay, T_{atm}, P_{atm}, GRS, PAR, ET_o, rain, w, and EVI. Correspondingly, we used artificial neural network (ANN1) and ANN2 for FCO_2 and ANN5 and ANN6 for FO_2 (Table 1). To develop and train the ANN architectures, we utilized the intelligent problem solver (IPS) tool in Statistica® v.7.0, to return the 10 ANNs with the lowest root mean square errors (RMSEs); the ANNs were randomly presented. The best model under each architecture was selected for each estimate (i.e., FCO_2 and FO_2). The execution of the training algorithm was stopped when the RMSE was less than 1% (Haykin, 2001).

Step 2: Although there is no defined number of input parameters in the construction of a model (Ozturk et al., 2011), reducing the number of inputs can contribute to a lower model complexity. Therefore, in step two, we performed a sensitivity analysis (SA) of the neural network. The IPS provides an SA methodology proposed by Hunter et al. (2000) to reduce the complexity of the problem. Flowchart of methodology used in this study (Fig. S2 in the Supplementary Material).

According to the procedure described, the SA used in this study iteratively rejected input variables and evaluated the network error. The measure of a variable's sensitivity is the ratio of the error without the variable to the original error, which provides the sensitivity score (Anagu et al., 2009; Hunter et al., 2000). The variables were ranked in order of importance based on the ratio.

From the sensitivity results for ANN1 and ANN3 for FCO_2 and ANN5 and ANN7 for FO_2, a new subset of data was formed, and a new training of multilayer perceptron (MLP) and radial basis function (RBF) was carried out to estimate FCO_2 and FO_2. Therefore, for FCO_2, the subset inputs of ANN2 consisted of 14 variables, namely P, ET_o, Cstock, T_s, Nstock, pH, H + Al, SB, macro, CN, sand, Bd, clay, and EVI while there were 17 variables for ANN4, namely SRG, EVI, P, Cstock, PAR, pH, H + Al, SB, micro, macro, CN, T_s, silt, clay, P_{atm}, ET_o, and sand (Figs. 1 and 2). For FO_2, the inputs of ANN6 were micro, EVI, silt, Cstock, macro, SWC, pH, H + Al, CN, SB, CTC, sand, and T_s; and those of ANN8 were SW, CEC, macro, sand, silt, clay, T_{atm}, P_{atm}, w, and CN (Fig. S3 and Fig. S4).

Adaptive neuro-diffuse FIS (ANFIS)

Jang (1993) developed a hybrid multilayer feedforward neural network and an FIS (ANFIS), which combines the advantages of both sources, making it the most powerful artificial intelligence technique (Kharb et al., 2014).The major advantage of the ANFIS model is its computational efficiency and adaptability (Ghadernejad et al., 2018). Hence, it can approximate any function. In this study, the Takagi and Sugeno fuzzy system is employed to formalize a systematic approach to generate fuzzy rules from a set of input and output data (Buragohain & Mahanta, 2008). The basic structure of an FIS is divided into fuzzification, a rule base, inferencing, and fuzzification.

A disadvantage of ANFIS is the restriction of the number of parameters defined in the model. If the number of inputs exceeds five layers, the training time takes too long to assign the appropriate values to the parameters of the membership function (Dehghani et al., 2019; Kaab et al., 2019). The most common membership functions are triangular, trapezoidal, and Gaussian (Klir & Yuan, 1995). In this study, the trapezoidal type is used (Riza et al., 2015). Although ANFIS has a high mapping ability of input to output similar to a black-box model, it suffers from a long training time that inhibits appropriate value assignments to the membership function parameters (the basis of fuzzy set theory).

Thus, several combinations of inputs were used for FCO_2 and FO_2. For FCO_2, these included pH, H + Al, ET_o, SRG, and PAR. For FO_2, these included T_s, SWC, P_{atm}, Nstock, and SB. In this study, the ANFIS was performed using R software (version 4.0.2; R Core Team, 2019, R Foundation for Statistical Computing, Vienna, Austria), using the R packages frbs: Fuzzy Rule-Based Systems for Classification and Regression Tasks (Riza et al., 2015).

Table 1 Classification of performance of statistical performance indicator (Camargo & Sentelhas, 1997)

			r	R^2 adj	RMSE	MAE	MAPE	*d*	Performance (*c*)		F	*p*-value
CO_2 emission in soil (µmol m^{-2} s^{-1})												
ANN1	MLP 17:17–9-1:1	Calibration	0.822	0.675	0.854	0.826	0.171	0.893	0.734	Good	493.3	2.20E − 16
		Validation	0.731	0.535	0.967	0.871	0.173	0.847	0.619	Median	135.5	2.20E − 16
Sensitivity analysis												
ANN2	MLP 14:14–14-10–1:1	Calibration	0.765	0.585	0.914	0.888	0.172	0.853	0.652	Median	344.2	2.20E − 16
		Validation	0.712	0.507	0.981	0.831	0.174	0.834	0.594	Sufferable	118.9	2.20E − 16
ANN3	RBF 17:17–68-1:1	Calibration	0.737	0.542	0.884	0.945	0.191	0.834	0.614	Median	284.6	
		Validation	0.708	0.501	0.872	0.880	0.183	0.817	0.578	Sufferable	117.3	*p*-value
Sensitivity analysis												
ANN4	RBF 17:17–25-1:1	Calibration	0.675	0.453	0.867	1.001	0.207	0.785	0.530	Sufferable	188.7	*p*-value
		Validation	0.645	0.417	0.862	1.036	0.207	0.758	0.492	Bad	85.29	1.38E − 15
O_2 influx (mg m^{-2} s^{-1})												
ANN5	MLP 14–10-1:1	Calibration	0.387	0.15	0.156	0.144	0.717	0.612	0.237	Insufficient	14.12	< 0.0001
		Validation	0.323	0.104	0.196	0.271	1.079	0.491	0.158	Insufficient	4.647	< 0.01
Sensitivity analysis												
ANN6	MLP 13–11-1:1	Calibration	0.769	0.591	0.142	0.142	0.516	0.855	0.658	Median	123.1	< 0.0001
		Validation	0.677	0.458	0.133	0.139	0.616	0.793	0.537	Sufferable	69.43	< 0.0001
ANN7	RBF 13–16-1:1	Calibration	0.452	0.204	0.093	0.109	0.607	0.646	0.292	Insufficient	21.3	< 0.0001
		Validation	0.309	0.095	0.114	0.281	0.69	0.346	0.107	Insufficient	6.71	< 0.01
Sensitivity analysis												
ANN8	RBF 10–3-1:1	Calibration	0.407	0.166	0.122	0.229	0.905	0.512	0.209	Insufficient	15.71	< 0.0001
		Validation	0.861	0.741	0.079	0.147	0.861	0.517	0.178	Insufficient	5.231	< 0.01
CO_2 emission in soil (µmol m^{-2} s^{-1})												
ANN9	ANFIS	Calibration	0.436	0.19	1.187	1.274	0.25857	0.666	0.291	Insufficient	776.8	< 0.01
		Validation	0.410	0.16	1.020	1.224	0.25714	0.637	0.261	Insufficient	385.8	< 0.01
O2 influx (mg m^{-2} s^{-1})												
ANN10	ANFIS	Calibration	0.385	0.14	0.1255	0.199	0.55487	0.462	0.178	Insufficient	6.414	< 0.01
		Validation	0.542	0.29	0.1100	0.134	0.81404	0.684	0.371	Insufficient	7.581	< 0.01

The sequence of numbers indicates the number of input variables, the number of neurons in the first layer, and the number of neurons in the second layer, respectively

r Pearson correlation coefficient: R^2 adj, *RMSE* root mean square error, *MAE* mean absolute error, *MAPE* mean absolute percentage error, *ANN* artificial neural network, *MLP* multilayer perceptron neural network, *ANFIS* adaptive neuro-fuzzy inference system, *RBF* radial basis function, agreement index, *d* Willmott's agreement index, *c* confidence coefficient, *p-value* probability of the occurrence of H0

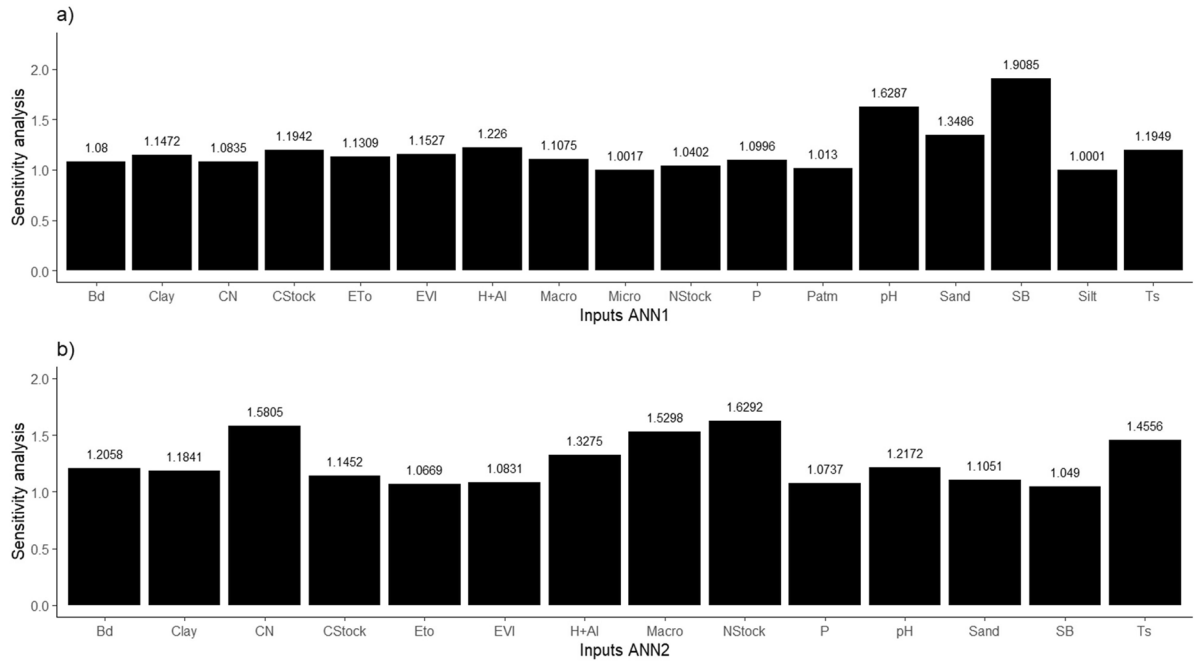

Fig. 1 Sensitivity analysis. Relative importance of ANN input variables in estimating soil CO_2 emission. Multilayer perceptron architecture. **a** ANN1 input datasets and **b** ANN2 input datasets

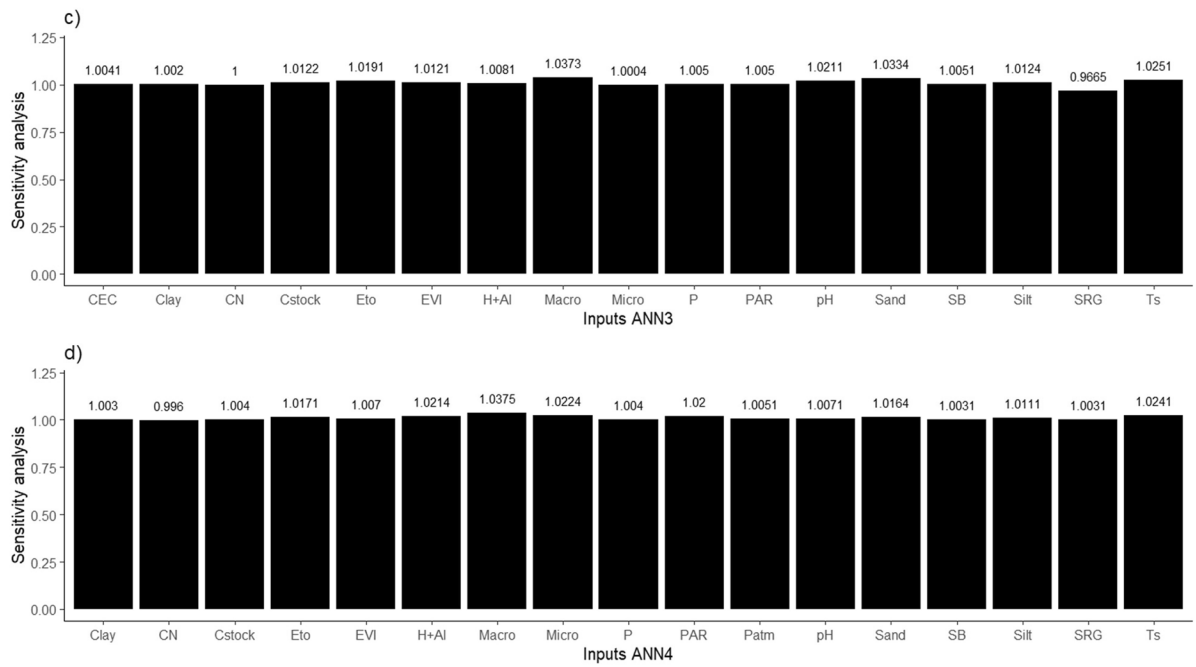

Fig. 2 Sensitivity analysis. Relative importance of ANN input variables in estimating soil CO_2 emission. Radial basis function (RBF) architecture. **c** ANN3 input datasets and **d** ANN4 input datasets

Criteria for model assessment

The evaluation of the efficiencies of the developed models, including accuracy and agreement, was assessed using statistical criteria of the Pearson correlation coefficient (r) (Eq. 1), mean absolute error (MEA) (Eq. 2), mean absolute percentage error (MAPE) (Eq. 3), agreement index (d) (Eq. 4) (Willmott, 1981), and confidence coefficient (c) (Eq. 5) (Camargo & Sentelhas, 1997), presented a scale varying from values smaller or equal to 0.40 (very bad performance) to values greater than 0.85 (excellent performance).

$$r = \frac{\left(N\sum_{i=1}^{N}\left(X\mathrm{obs}_1 Y\mathrm{obs}_1\right) - \sum_{i=1}^{N} X\mathrm{obs}_1 \sum_{i=1}^{N} Y\mathrm{obs}_1\right)\left(\left(N\sum_{i=1}^{N} Y\mathrm{obs}_1\right)^2 - \left(\sum_{i=1}^{N} Y\mathrm{obs}_1\right)^2\right)}{\sqrt{\left(\left(N\sum_{i=1}^{N} X\mathrm{obs}_1\right)^2 - \left(\sum_{i=1}^{N} X\mathrm{obs}_1\right)^2\right)}}, \quad (1)$$

$$\mathrm{MEA} = \frac{\sum_{i=1}^{N}(Y\mathrm{obs}_1 - Y\mathrm{est}_1)}{N}, \quad (2)$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\left|(A_i - P_i)\right|}{A_i}100\right), \quad (3)$$

$$d = 1 - \frac{\sum_{i=1}^{N}\left(Y\mathrm{obs}_1 - Y\mathrm{est}_1\right)^2}{\sum_{i=1}^{N}\left(\left|Y\mathrm{est}_1 - \overline{\mathrm{Y}}\right| + \left|Y\mathrm{obs}_1 - \overline{\mathrm{Y}}\right|\right)^2}, \quad (4)$$

$$c = c.d, \quad (5)$$

where n is number of data points, $X\mathrm{obs}_1$ is the observed value of X, $Y\mathrm{obs}_1$ is the observed value of Y, $Y\mathrm{est}_1$ is the estimated value of Y, and $\bar{Y}$ is the observed mean of Y.

Results and discussion

Soil CO_2 emission (FCO_2) and influx of O_2 into soil (FO_2)

All the models generated by the ANN and adaptive neuro-fuzzy inference system (ANFIS) for FCO_2 and FO_2 were significant ($p < 0.01$) (Table 1). In this study, the best performance was observed for FCO_2. The R^2 value ranged from 0.45 to 0.67 and from 0.41 to 0.50, respectively, after calibration and validation. During calibration, the values of FO_2 ranged from 0.15 to 0.59; during validation, they ranged from 0.095 to 0.74. Both phases had high associated perceptual errors. Neural network architecture is a key factor that influences learning time and network performance. The two models MLP and RBF that were used in the study were compared. The results of the neural network estimates used for FCO_2 and FO_2 prediction are presented in Table 1 and Figs. 3 and 4.

Generally, for the FCO_2 estimation, the evaluation criteria used for calibration and validation, namely RMSE and mean absolute error (MAE), indicated how close the data points were to the line of best fit. According to Table 1, all the models were within the acceptable range; a value of zero indicates a perfect fit to the data.

For FCO_2, the RMSE, MAE, and values were within the acceptable range, as mentioned in Table 1. During calibration, they were below one. The RMSE values ranged from 0.85 to 0.98 and the MAE values ranged from 0.83 to 0.89 in both the calibration and validation phases for the MLP models. For the RBF architecture, the values varied between RMSE = 0.86 to 0.88 and MAE = 0.88 to 1.04.

The mean absolute percentage error (MAPE) indicated that the MLP architecture obtained a better fit in relation to that obtained by the RBF network, varying around 17%, whereas RBF varied between 18 and 20%. Regarding precision, the coefficient (r) between the actual data and predicted data of the FCO_2 for the MLP and RBF ANNs was between 0.645 and 0.822, demonstrating a good relationship between the estimated and observed data.

Our results revealed that after both phases of ANN processing (after SA), the performances of the MLP models were superior to those of the RBF model, and the results were satisfactory for all the datasets. According to Hamedi and Dehdashti

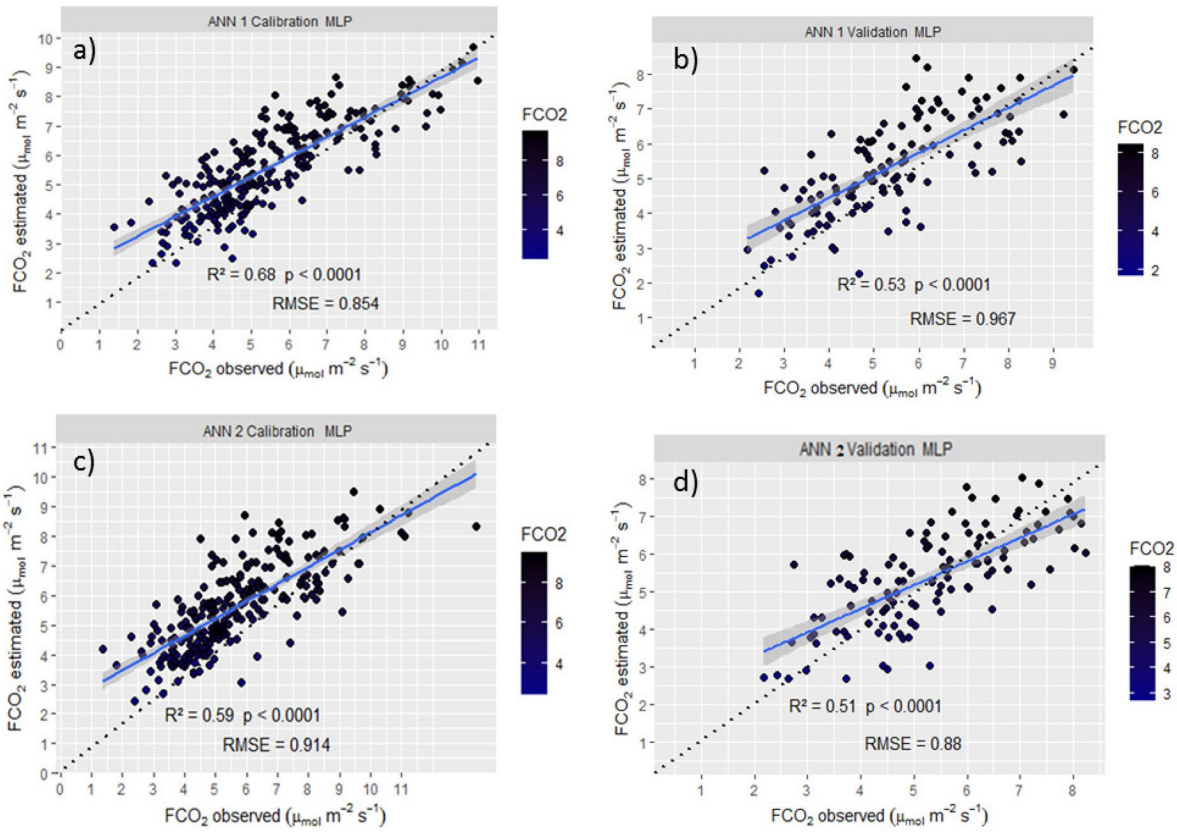

Fig. 3 Performance of soil CO_2 emission (µmol m^{-2} $s.^{-1}$) during the calibration and validation phases of artificial neural networks (ANNs) multilayer perceptron. **a** Calibration ANN1 and **b** validation ANN1; ANN2 after sensitivity analysis, **c** calibration ANN2 and **d** validation ANN2

Jahromi (2021), the smallest errors of calibration and validation and the maximum of the coefficient of determination (R^2) need to be considered when choosing the best network parameters, which shows the ability of the model in explaining the data observed in scatter plots and determining the appropriateness of the network size. ANN1 represented the best MLP architecture and demonstrated that 53% of the temporal variability of FCO_2 could be explained with 17 input variables comprising soil physical–chemical parameters, agro-meteorological parameters, and enhanced vegetation index (EVI) (Table 1). We believe that this result was adequate to predict the non-linear and complex relationship between FCO_2 and all the environmental variables that have been considered in this study.

After the SA was carried out by the IPS represented by ANN2, the exclusion of three variables, namely silt, P_{atm}, and micro in the network dataset resulted in the deterioration of the performance of the model. The coefficient of determination in the ANN2 model was less than that of the ANN1 model under both the calibration and validation phases, with values of 0.58 and 0.50, respectively. This result indicated that despite the low contribution value of these variables, they influenced the temporal variability of soil CO_2 emissions.

As observed in this study, RBF networks generally required more neurons and tended to use more parameters from the learning dataset because they had a single hidden layer (Boniecki et al., 2020). Models ANN3 and ANN4 had 17 input variables, even after SA.

The concordance index, which relates the distance of the observed values from the estimated values of the standard model, showed that during

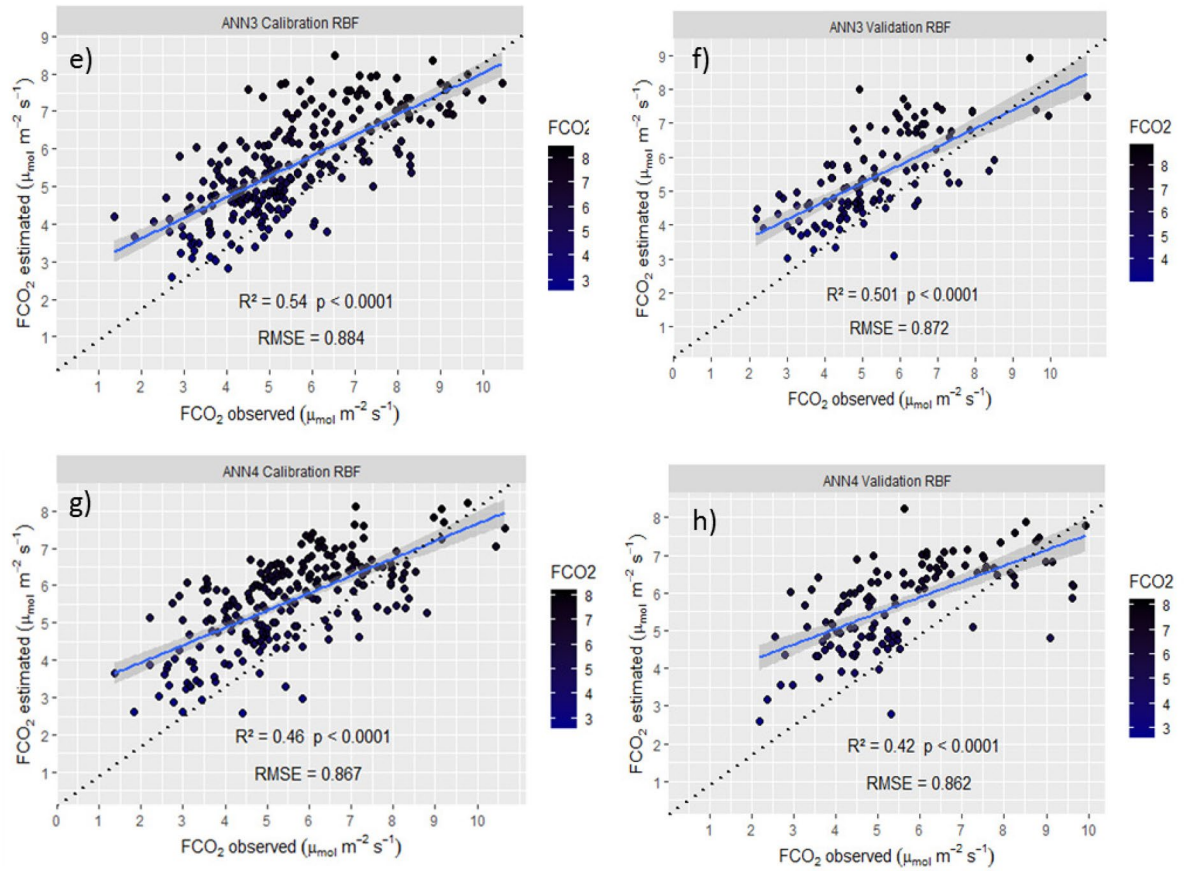

Fig. 4 Performance of soil CO_2 emission (μmol m^{-2} $s.^{-1}$) during the calibration and validation phases of artificial neural networks (ANNs) radial basis function (RBF). **e** Calibration ANN3 and **f** validation ANN3; ANN4 after sensitivity analysis, **g** calibration ANN4 and **h** validation ANN4

calibration, the values were 0.83 to 0.89 for MLP and 0.75 to 0.83 for RBF. These results corroborated with the good classification performance according to the classification index (Camargo & Sentelhas, 1997) for both ANNs. Regarding the validation phase, the MLP presented a bad calibration ($d = 0.75$), whereas the RBF presented a good one ($d = 0.84$).

We believe that the superior predictive ability of the MLP over that of the RBF in obtaining the same results is related to the presence of smaller hidden neurons, since RBF follows hidden transformation for the output space and a linear combination in contrast to the approach of the MLP hidden layer, which executes a set of algorithms to compute the non-linear transformation of the input data. Owing to the complexity involved in estimating the FCO_2, our results indicated that the accuracy of artificial neural networks (ANNs) was related to the increase in the number of input variables (Schaap & Leij, 1998).

For FO_2, all models showed inferior performance compared with those of FCO_2. Despite the low RMSE values for calibration (0.122 to 0.144) and validation ranged from 0.079 to 0.196, there are high percentage errors associated with the area at FO_2, ranging from 51 to 90% in calibration and 61 to 86% in validation, indicating a lower accuracy. We believe that the size of the dataset for neural-network training ($n = 176$) affected the prediction ability of the ANNs.

The FCO_2 dataset is compact ($n = 500$). In an ANN, generalization, or the ability to respond appropriately to previously unseen patterns, is influenced primarily by the size of the training set and its representativeness

with respect to estimation, problem complexity, and ANN architecture (Haykin, 2001). Additionally, the observed FO_2 values in the field are very low. In the original dataset, over the 193 study days, the values ranged from 0.030 to 2.02 (mg m^{-2} s^{-1}, with a mean value of 0.371 mg m^{-2} s^{-1}), and in the soil, the temporal variability of oxygen mainly related to the frequency of precipitation events (Liptzin et al., 2011). The diffusion of O_2 in soil is conditioned by its physical properties (Neira et al., 2015). MAPE is often used in model performance evaluation because it is easy to understand (de Myttenaere et al., 2016).

However, the very small values close to zero observed in the dataset present a disadvantage, because in the calculation, MAPE divides the absolute error by the actual data. Thus, values close to zero can significantly increase the MAPE. We believe that this may have caused the values observed when training the ANNs of FO_2. Like the FCO_2 ANN, after reducing the dataset with SA, there was an improvement in the correlation of the estimates. For FO_2, the estimation using MLP performed slightly better than the RBF. However, with the exception of ANN 6, all ANNs had classification index (c) considered as "very bad or insufficient," below 0.40 in the calibration and validation phases (Camargo & Sentelhas, 1997).

When comparing the performance of the ANNs to ANFIS (Table 1; Figs. 3, 4, and 5), the RMSE, MEA, and MAPE values indicate an inferior performance of the ANFIS. In the ANFIS, the agreement index value in the FCO_2 estimate ranged from 0.63 to 0.66 for the calibration and validation periods, indicating a low relationship between observed and modeled values. The same was found for FO_2, with values ranging from 0.46 to 0.68. In both estimates, the phases (i.e., calibration and validation) were classified as "bad" according to the classification index (Camargo & Sentelhas, 1997).

The ANFIS obtained RMSE and MEA errors higher than those of MLP and RBF. For the FCO_2 ANN, the RMSE values ranged from 1 to 1.18 μmol m^{-2} s^{-1} and MEA from 1.22 to 127 μmol m^{-2} s^{-1}. These results indicate low accuracy with coefficient of determination values equal to 0.19 and 0.16 for calibration and validation phases, respectively. For FO_2, the performance was also poor in both validation and calibration. Despite the low RMSEs (0.11 to 1.12 mg m^{-2} s^{-1}) and MEAs (0.13 to 0.19 mg m^{-2} s^{-1}), validation and calibration, respectively, those observed for the FO_2 ANN had high MAPE values, reflecting a low accuracy of 55% for calibration and 88% for validation.

We expected a better performance for ANFIS compared with the two ANNs, owing to the higher degree of robustness and fault tolerance provided by the model. Owing to the complexity of FCO_2 and FO_2 modeling, the limitation regarding the number of system inputs compromised the estimation performance.

Overall, it is interesting that in most of the models studied, soil temperature and macroporosity formed part of the dataset in the calibration of the models, except for ANFIS, which only considered soil temperature in the estimation of FO_2 and macroporosity, and did not contribute to any of the estimates (Figs. 1 and 2 and Fig. S3, Fig. S4).

Macroporosity is related to the diffusion of CO_2 and O_2 gases into the soil, and the soil temperature is a determining factor in the decay of organic matter (Curiel Yusteet al., 2007), with high temperatures accelerating microbial and root activity (Silva-Olaya et al., 2013), except in tropical soils, since the soil already has a temperature in ideal conditions for favoring biological activity (Adachi et al., 2006), and was not a limiting factor for the estimation in the ANFIs model.

Although several studies have reported temperature and moisture as the controlling factors on the temporal dynamics of soil respiration (Davidson et al., 1998; Lal, 2009), and directly influence oxygen consumption, either by the activity of root respiration and microorganisms that in their metabolic processes that consume soil oxygen or due to oxygen transport in soils occurs mainly by diffusion (Neira et al., 2015). In this study, soil water content did not contribute to the dataset during model calibration. This observation may be related to the covariation that exists between soil temperature, evapotranspiration, and precipitation, inhibiting the determination of the individual effects of each variable (Wu et al., 2014; Davidson et al., 1995), in this study, soil water content did not contribute to the dataset during model calibration.

It is worth noting that although ANFIS did not incorporate soil temperature in its calibration dataset to FCO_2; solar radiation played an important role in this estimation. Solar radiation and air temperature are among the main factors influencing the

variation in soil temperature (Reichardt & Timm, 2012), in addition to soil water content and texture, and the surface area covered by leaf litter and plant canopies (Paul et al., 2004). Moreover, ANFIS was the only model in which evapotranspiration played an important role in the calibration dataset for the estimation of FCO_2.

It is worth noting that, in addition to the physiological process of plant transpiration, evapotranspiration is the result of physical processes such as energy exchange between the earth's surface and the atmosphere, and is related to solar radiation, soil water content, precipitation, and wind (Liang et al., 2010). We also observed that only ANFIS was sensitive to the Nstock in the model calibration dataset for the estimation of FO_2. In the soil, heterotrophic, anaerobic, and aerobic microorganisms that consume oxygen are responsible for the mineralization of organic N,

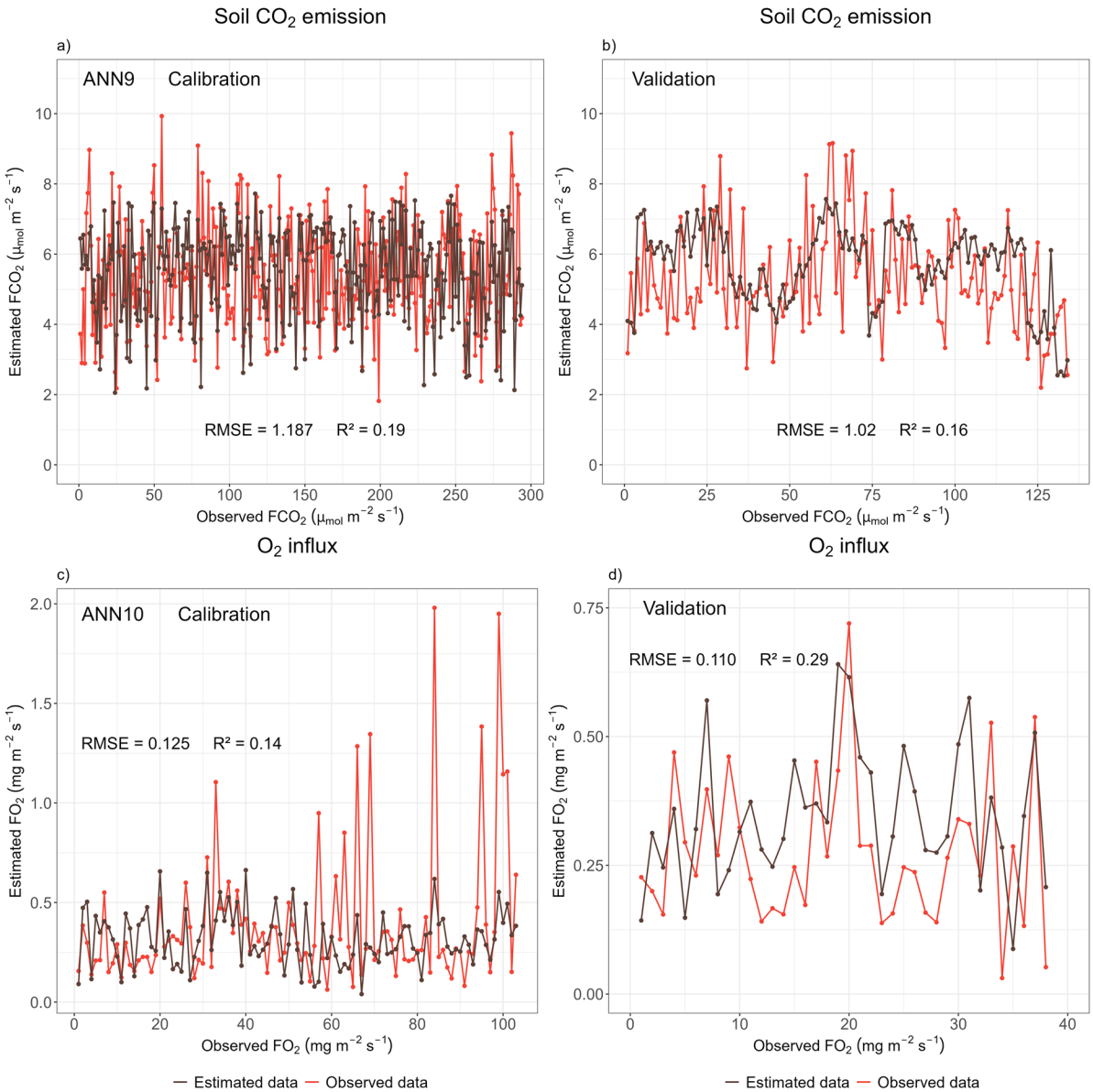

Fig. 5 Model soil CO_2 emission (μmol m^{-2} s^{-1}) and O_2 influx (mg O_2 m^{-2} s^{-1}) during the calibration and validation phases of the adaptive neuro-diffuse inference system. Black line: observed data; red line: estimated. **a** ANN9-FCO_2 calibration, **b** ANN9-FCO_2 validation, **c** ANN10-FO_2 calibration, **d** ANN10-FO_2 validation

which constitutes the process by which N becomes available to plants. Thus, although the Cstock did not contribute to the calibration of the dataset, it is known that soil organic matter is closely related to the amount of soil N, and temporal changes in the soil C and Nstocks of reforested areas tend to show similar temporal patterns (Li et al., 2021). In addition, during the calibration of the ANFIS model for FO_2 estimation, atmospheric pressure was observed to be a highly sensitive model parameter. The entry of oxygen into the soil occurs by diffusion, as described by Fick's law, is pressure dependent (Neira et al., 2015).

Studies conducted on various ecosystems and soils demonstrated good correlations for FCO_2 estimation using physicochemical properties, for example, for your study (Freitas et al., 2018) observed good MLP results with a backpropagation algorithm for a sugarcane area situated in Brazil. The R^2 of the model was 0.91, with a MAPE error of 18%. For ANFIS, (Canteral, 2020) determined an R^2 value ranging from 0.53 to 0.71 and a MAPE error of 28.84–50.46% for tropical forest soils. However, soil is a complex and spatially heterogeneous mixture of minerals, carbon stock, roots, and microorganisms, and climatic factors such as precipitation and radiation contribute to the formation of different microclimates that drive temporal variability in soil respiration (Rubio & Detto, 2017). Therefore, comparing our results with other estimates does not necessarily provide insight into the dynamics of respiration in our ecosystem. Regarding oxygen consumption, thus far, we have not observed estimates with machine learning for tropical soils.

Principal component analysis

We performed a PCA to aid in the temporal dynamic compression of FCO_2 and FO_2 and identify the main factors involved in this process. The first three dimensions of the analysis expressed 53.4% of the total dataset inertia, indicating that 53.4% of the individual (or variable) cloud total variability was explained by the plane. Generally, the best results are obtained when the original variables are highly or inversely correlated (Manly & Alberto, 2008). Moreover, the first component was the most important of the 10 components (Fig. 6).

However, since FCO_2 and FO_2 contributed more to the first two components, we have only discussed the linear combinations of the variables in these two components (Table 2 and Fig. 7).

As shown in Table 2, the variables that contributed the most to PC1 in order of relevance were represented by SB (0.94), pH (0.93), Nstock (0.87), macro (0.59), Cstock (0.55), clay (0.47), *P* (0.40), CEC (0.34), FO_2 (0.26), inversely correlated H + Al (−0.87), Bd (−0.74), silt (−0.69), micro (−0.61), sand (−0.24), and CN (−0.23). This selection of variables indicated that this process might be related to heterotrophic soil respiration, derived from the decomposition of soil organic matter, influenced by substrate quality, and controlled by soil physical characteristics.

For PC2, the following variables were retained in the following order: ET_o (0.89), T_{air} (0.87), GSR (0.83), PAR (0.75), T_s (0.70), EVI (0.67), and FCO_2

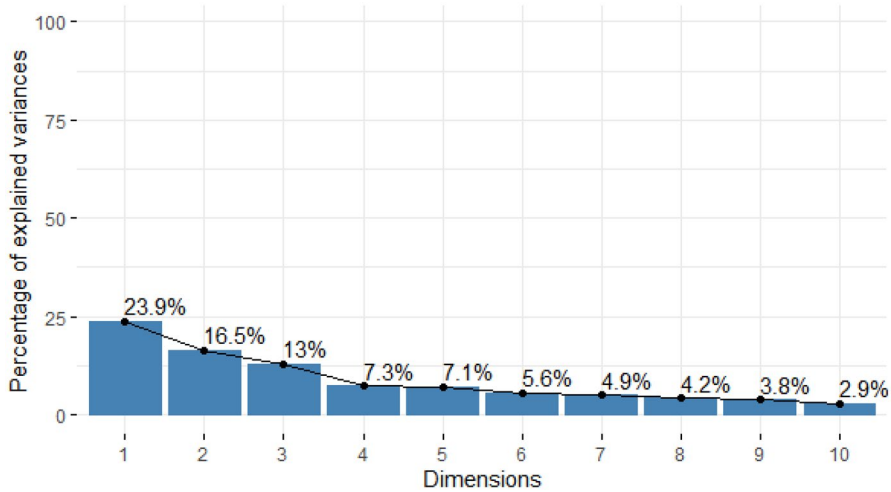

Fig. 6 Decomposition of total inertia

Table 2 Correlation between FCO_2 and FO_2 with environmental variables

Principal component	PC1	PC2	PC3
Explained variance (%)	23.9	16.5	13
Correlation			
FCO_2	0.10	0.51	−0.14
FO_2	0.26	0.14	−0.02
Ts	−0.05	0.70	−0.01
SWC	0.02	0.09	−0.11
P	0.40	−0.04	0.58
pH	0.93	0.01	0.01
H + Al	−0.87	−0.04	0.30
SB	0.94	0.02	0.08
CEC	0.34	−0.06	0.62
Macro	0.59	0.03	−0.37
Micro	−0.61	0.02	−0.02
Bd	−0.74	−0.08	−0.16
Cstock	0.55	−0.03	0.82
Nstock	0.86	0.01	−0.10
CN	−0.24	−0.05	0.84
Sand	−0.23	−0.01	0.71
Silt	−0.69	−0.07	−0.16
Clay	0.47	0.04	−0.55
T_{air}	−0.05	0.87	0.06
P_{atm}	0.04	−0.53	−0.06
GSR	−0.05	0.83	0.06
PAR	−0.04	0.75	0.06
ET_o	−0.05	0.89	0.06
Rain	0.01	−0.04	−0.02
W	0.01	−0.20	0.10
EVI	−0.01	0.67	0.01

FCO_2 soil CO_2 emission, *FO_2* O_2 influx, *Ts (°C)* soil temperature, *SWC* soil water content, *P (mg dm^{-3})* soil *P* content, *pH in CaCl2* hydrogen potential, *H + Al* potential acidity, sum of bases (K+, Ca2+, and Mg2+), *CEC (mmolc dm^{-3})* cation exchange capacity, *macro (m^3 m^{-3})* macroporosity, *micro (m^3 m^{-3})* microporosity, *Bd* bulk density, (g dm^{-3}), *Cstock* (mg ha − 1), carbon stocks (mg ha − 1), *Nstock (mg ha − 1)* nitrogen stocks, CN ratio, *sand* sand content (g kg^{-1}), *silt* silt content (g kg^{-1}), *clay* clay content (g kg^{-1}), *T_{atm}* temperature atmosphere, *P_{atm} (kPa)* atmospheric pressure, *GSR (MJ m^{-2} day^{-1})* global solar radiation, PAR (µmol m^{-2}), photosynthetically active radiation, *(ET_o) EVI* enhanced vegetation index, rain (mm), *w* wind speed *w* (m s^{-1}), *EVI* enhanced vegetation index

(0.51), and inversely correlated P_{atm} (−0.53), *W* (−0.20), and FO_2 (0.14). Although the variables retained in this component indicate a greater contribution of root respiration to total respiration, it is very difficult to separate them in the field, as both occur in association with and compete for oxygen availability (Ben-Noah & Friedman, 2018). It is known that the atmospheric conditions of the region, weather factors, and the type of soil cover interfere with the dynamics of FCO_2 and FO_2 (De Lucena et al., 2023).

Leaf litter is the primary source of organic ecosystems. The organic compounds in these plant residues are composed of carbohydrates with different degrees of complexity as well as lignins and polyphenols, which are resistant to decomposition. The quality of these molecules is expressed by the CN ratio, and both the quality and quantity of the residues influence the dynamics of dissolved organic carbon and nitrogen in the soil (Zhou et al., 2015).

When biochemical processes such as the decomposition of litter occur in well-aerated soil, microorganisms consume oxygen and CO_2 are released. Moreover, nutrients such as phosphorus, nitrogen, and sulfur are immobilized or released, which is fundamental for the recycling of soil nutrients (Moreira & Siqueira, 2006). This was revealed by the direct relationship between SB, CEC, *P*, and pH with FCO_2 (Fig. 7). Among the abiotic factors affecting soil microorganisms, pH is an important factor in the formation, size, and distribution of soil microbial diversity (Cao et al., 2017; Liu et al., 2014), and also influences nutrient availability. In our study, the measured pH value was 4.4, which is within the expected range for this soil (Moreira & Siqueira, 2006).

In addition, during decomposition, rearrangements and formation of material that are unaffected by microbial action may occur because of the dynamics of soil aggregates. This process alter the storage of C and N in the soil and in the humic fractions of organic matter (Golchin et al., 1994; Silva et al., 2016), reflecting the temporal dynamics of FO_2 and FCO_2.

In the ecosystem analyzed in this study, the temporal dynamics of FO_2 and FCO_2 are more complex when compared with monoculture reforestation (Li et al., 2012), mainly due to phenological differences. The mixed forest examined in this study, constituting various colonizing plants and 21 species of trees with deciduous or subcaducifolic species consisting of pioneer, secondary, and late species, is characterized by substrates with different degrees of lability and seasonal variations in plant residues (Davidson et al., 2006). The CN ratio of the study area revealed an indirect association with C and Nstocks, as well as FCO_2 and FO_2, indicating that the vegetation type significantly influences the temporal dynamics of soil respiration.

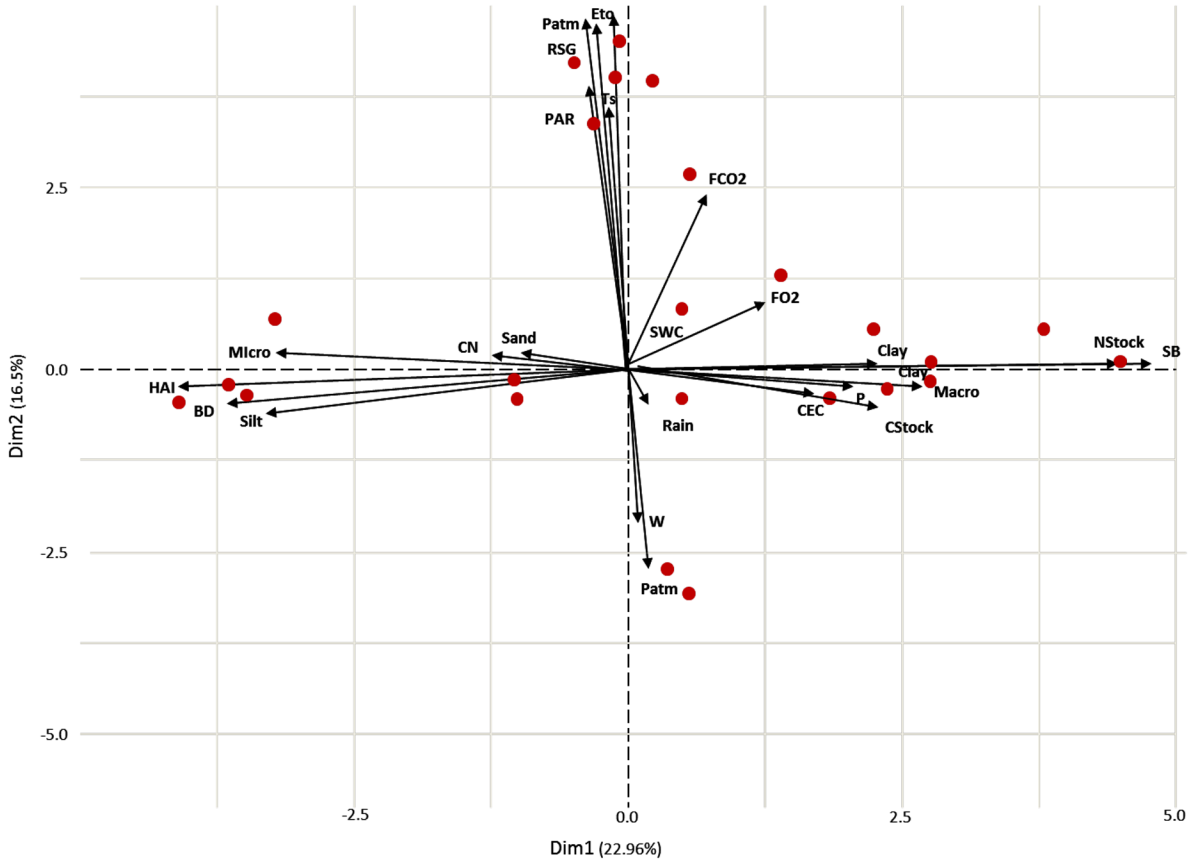

Fig. 7 Biplot graph of environmental variables and soil characteristic

Physical factors influence both the quality of the residue and the transport of gases in the soil (Brady & Weil, 2013). As mentioned previously, based on our observations, we believe that soil temperature was not a limiting factor in this study. Moreover, studies on forest ecosystems suggest that autotrophic respiration is more sensitive to temperature than heterotrophic respiration (Chen et al., 2017; Mäkiranta et al., 2008). Therefore, we did not consider the PC2 results anomalous since the existing linear correlations in this PC indicate that FCO_2 is most likely a by-product of root respiration. Moreover, considerable evidence exists to support the correlation between the physiological processes of roots and photosynthetic activity (Tang et al., 2005; Kuzyakov & Gavrichkova, 2010; Wang et al., 2010).

In addition, these studies also suggest significant correlations between the hyperspectral vegetation indices, air temperature, and soil respiration, for which 82% of the temporal variability of respiration can be attributed in broadleaved evergreen and mature broadleaved deciduous forests (Wu et al., 2014). The low correlation between the water content and both components can be explained by either covariation with ET_o or soil temperature (Davidson et al., 1998; Schwendenmann et al., 2003).

We believe that the high negative contribution of microporosity to PC1 is related to soil compaction. Note that both microporosity and soil density have direct and opposite associations with macroporosity. The physical characterization of the area described in the methodology indicates a density value of 1.34 g dm^{-3}, which is higher than the value expected for clayey soils (Arshad et al., 1997). It is known that soil density influences the total porosity and distribution of pore sizes, influencing water retention in the soil and hindering the process of gas transport in the soil by consuming oxygen and CO_2 output from the

soil into the atmosphere (Bicalho et al., 2014; Teixeira et al., 2012; Tormena et al., 1998).

The biplot Fig. 7 indicates substantiating information. Note that the vectors and dispersion points of SWC, FO_2, FCO_2, clay, C, and Nstock participate in Group I, located on the right side of the biplot. In tropical soils, latosols tend to be rich in clays, which are directly linked to water availability; thus, water availability controls the oxygen dynamics of these soils and may lead to periods when the O_2 demand exceeds supply, influencing the dynamics of microbial community activity (LIptzin et al., 2011). A possible group (Group II), formed to the left of the biplot would be predominantly related to the dynamics of soil gas transport, except for the CN ratio.

Conclusion

In this study, we obtained the best estimate of CO_2 emission from the soil using the multilayer perceptron architecture neural network model, thus explaining 53% of the temporal variability of soil respiration in a reforested area with native plants.

After sensitivity analysis, the multilayer perceptron exhibited a loss in explaining the temporal variability of CO_2 emission, when three variables from the model were excluded. The best explanation of the estimation with radial-basis function architecture was found in the architecture with a higher number of neurons in the intermediate layer.

In contrast to the probabilities, the size of the soil oxygen influx database was insufficient for the network to generalize the neural network and adaptive neuro-fuzzy inference system models owing to the complex relationships between CO_2 emission and O_2 influx in mixed species reforested soil. Therefore, the results of this study serve as an incentive for the future research.

We obtained an inferior performance for the adaptive neuro-fuzzy inference system compared with the neural networks in both the flux estimates. The input variable defined for each model significantly varied when the variables were combined. Soil temperature and macroporosity were found to be good predictors of soil CO_2 emission. The adaptive neuro-fuzzy inference system model was very sensitive to Nstock and atmospheric pressure in calibrating and estimating soil CO_2 emission and O_2 influx, respectively.

Principal component analysis showed associations between heterotrophic and autotrophic respiration. FCO_2 was best correlated with variables that are linked to root respiration, while FO_2 correlated with variables that are linked to heterotrophic respiration. The quality of litters, which was expressed with CN ratio, is one of the factors that control this dynamics. Given the complexity of the data obtained in this study, neural networks outperformed ANFIs in both estimates (FCO_2 and FO_2). New studies and data mining analytics need to be tested. ANNs have the potential to generate predictive FCO_2 and FO_2 models.

Author contribution Maria Elisa Vicentini: writing—original draft, investigation, methodology, and formal analysis. Paulo Alexandre da Silva: investigation and methodology. Kleve Freddy Ferreira Canteral and Wanderson Benerval De Lucena: investigation, methodology, and prepared figures. Mário Luiz Teixeira de Moraes, Rafael Montanari, and Marcelo Carvalho Minhoto Teixeira Filho: resources and methodology. Nelson José Peruzzi: review and editing. Newton La Scala Jr: conceptualization and supervision. Glauco De Souza Rolim: conceptualization and supervision. Alan Rodrigo Panosso: conceptualization, supervision, review and editing. All authors reviewed the manuscript.

Funding We are grateful to Fundação de Amparo à Pesquisa do Estado de São Paulo0–FAPESP (process no 2016/03861–5), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for support (Finance Code 001).

Data availability The data that support the findings of this study are available from the corresponding author (Vicentini, M.E.) upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Ethical approval All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors and are aware that with minor exceptions, no changes can be made to authorship once the paper is submitted.

Conflict of interest The authors declare no competing interests.

References

Adachi, M., Bekku, Y. S., Rashidah, W., Okuda, T., & Koizumi, H. (2006). Differences in soil respiration between different

tropical ecosystems. *Applied Soil Ecology*. https://doi.org/10.1016/j.apsoil.2006.01.006

Anagu, I., Ingwersen, J., Utermann, J., & Streck, T. (2009). Estimation of heavy metal sorption in German soils using artificial neural networks. *Geoderma, 152*(1–2), 104–112. https://doi.org/10.1016/j.geoderma.2009.06.004

Arshad, M. A., Lowery, B., & Grossman, B. (1997). Physical tests for monitoring soil quality. *Methods for Assessing Soil Quality, 49*, 123–141. https://doi.org/10.2136/sssaspecpub49.c7

Barlow, J., França, F., Gardner, T. A., Hicks, C. C., Lennox, G. D., Berenguer, E., & Graham, N. A. (2018). The future of hyperdiverse tropical ecosystems. *Nature*, *559*(7715), 517–526. https://doi.org/10.1038/s41586-018-0301-1

Bataglia, O. C., Furlani, A. M. C., & Teixeira, J. P. F. (1983). Methods of chemical analysis of plants = Métodos de Análise Química de Plantas. Instituto Agronômico de Campinas, Campinas, SP, Brazil. *Boletim Técnico, 78*.

Ben-Noah, I., & Friedman, S. P. (2018). Review and evaluation of root respiration and of natural and agricultural processes of soil aeration. *Vadose Zone Journal, 17*(1), 1–47. https://doi.org/10.2136/vzj2017.06.0119

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 CO2 emission and soil attributes in a sugarcane area. *Agriculture, Ecosystems & Environment, 189*, 206–215. https://doi.org/10.1016/j.agee.2014.03.043

Bond-Lamberty, B. (2018). New techniques and data for understanding the global soil respiration flux. *Earth's Future, 6*(9), 1176–1180. https://doi.org/10.1029/2018EF000866

Boniecki, P., Zaborowicz, M., Pilarska, A., & Piekarska-Boniecka, H. (2020). Identification process of selected graphic features apple tree pests by neural models type MLP. *RBF and DNN. Agriculture, 10*(6), 218. https://doi.org/10.3390/agriculture10060218

Brady, N. C., & Weil, R. R. (2013). *Elementos da natureza e propriedades dos solos* (3a ed., p. 704). Bookman: Porto Alegre.

Buragohain, M., & Mahanta, C. (2008). A novel approach for ANFIS modelling based on full factorial design. *Applied Soft Computing, 8*(1), 609–625. https://doi.org/10.1016/j.asoc.2007.03.010

Camargo, A. P., & Sentelhas, P. C. (1997). Avaliação do desempenho de diferentes métodos de estimativa da evapotranspiração potencial no estado de São Paulo. *Revista Brasileira de Agrometeorologia, 5*(1), 89–97.

Canteral, K. F. F. (2020). Aprendizado de máquina na modelagem temporal da emissão de CO2 do solo em áreas agrícolas no Cerrado brasileiro.

Canteral, K. F. F., Vicentini, M. E., de Lucena, W. B., de Moraes, M. L. T., Montanari, R., Ferraudo, A. S., & Panosso, A. R. (2023). Machine learning for prediction of soil CO2 emission in tropical forests in the Brazilian Cerrado. *Environmental Science and Pollution Research*, 1–20. https://doi.org/10.1007/s11356-023-26824-6

Cao, C., Zhang, Y., Qian, W., Liang, C., Wang, C., & Tao, S. (2017). Land-use changes influence soil bacterial communities in a meadow grassland in Northeast China. *Solid Earth, 8*(5), 1119–1129. https://doi.org/10.5194/se-8-1119-201

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment, 91*(3–4), 332–344. https://doi.org/10.1016/j.rse.2004.03.014

Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., & Wang, K. (2017). Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy*, *195*, 659–670. https://doi.org/10.1016/j.apenergy.2017.03.034

Curiel Yuste, J., Baldocchi, D. D., Gershenson, A., Goldstein, A., Misson, L., & Wong, S. (2007). Microbial soil respiration and its dependency on carbon inputs, soil temperature and moisture. *Global Change Biology, 13*(9), 2018–2035. https://doi.org/10.1111/j.1365-2486.2007.01415.x

Davidson, E. A. (1995). Spatial covariation of soil organic carbon, clay content, and drainage class at a regional scale. *Landscape Ecology, 10*, 349–362. https://doi.org/10.1007/BF00130212

Davidson, E. A., Belk, E., & Boone, R. D. (1998). Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global Change Biology, 4*(2), 217–227. https://doi.org/10.1046/j.1365-2486.1998.00128.x

Davidson, E. A., Janssens, I. A., & Luo, Y. (2006). On the variability of respiration in terrestrial ecosystems: Moving beyond Q10. *Global Change Biology, 12*(2), 154–164. https://doi.org/10.1111/j.1365-2486.2005.01065.x

de Almeida, R. F., de Bortoli Teixeira, D., Montanari, R., Bolonhezi, A. C., Teixeira, E. B., Moitinho, M. R., & Júnior, N. L. S. (2018). Ratio of CO2 and O2 as index for categorising soil biological activity in sugarcane areas under contrasting straw management regimes. *Soil Research*, *56*(4), 373–381. https://doi.org/10.1071/SR16344

de Araújo Santos, G. A., Moitinho, M. R., de Oliveira Silva, B., Xavier, C. V., Teixeira, D. D. B., Corá, J. E., & Júnior, N. L. S. (2019). Effects of long-term no- tillage systems with different succession cropping strategies on the variation of soil CO2 emission. *Science of the Total Environment, 686*, 413–424. https://doi.org/10.1016/j.scitotenv.2019.05.398

de Lucena, W. B., Vicentini, M. E., Santos, G. A. D. A., Silva, B. D. O., Da Costa, D. V. M., Canteral, K. F. F., & La Scala Jr, N. (2023). Temporal variability of the CO2 emission and the O2 influx in a tropical soil in contrasting coverage conditions. *Journal of South American Earth Sciences*, *121*, 104120. https://doi.org/10.1016/j.jsames.2022.104120

de Myttenaere, A. (2016). Offline evaluation of a predictive model: Application to recommendation algorithms and to mean absolute percentage error. HAL, 2016. http://gdmltest.u-ga.fr/item/tel-01395290/

de Oliveira Silva, B., Moitinho, M. R., de Araujo Santos, G. A., Teixeira, D. D. B., Fernandes, C., & La Scala Jr, N. (2019). Soil CO_2 emission and short-term soil pore class distribution after tillage operations. *Soil and Tillage Research, 186*, 224–232. https://doi.org/10.1016/j.still.2018.10.019

de Souza, L. C., & Procópio, L. (2021). The profile of the soil microbiota in the Cerrado is influenced by land use. *Applied Microbiology and Biotechnology, 105*(11), 4791–4803. https://doi.org/10.1007/s00253-021-11377-w

Dehghani, M., Riahi-Madvar, H., Hooshyaripor, F., Mosavi, A., Shamshirband, S., Zavadskas, E. K., & Chau, K. W.

(2019). Prediction of hydropower generation using grey wolf optimization adaptive neuro-fuzzy inference system. *Energies, 12*(2), 289. https://doi.org/10.3390/en12020289

dos Santos, H. G., Jacomine, P., dos Anjos, L. H. C., de Oliveira, V. A., Lumbreras, J., Coelho, M., ... & de Araujo Filho, J. C. (2022). Proposta de atualização da quinta edição do Sistema Brasileiro de Classificação de Solos: ano 2022.

Ebrahimi, M., Sarikhani, M. R., Sinegani, A. A. S., Ahmadi, A., & Keesstra, S. (2019). Estimating the soil respiration under different land uses using artificial neural network and linear regression models. *CATENA, 174*, 371–382. https://doi.org/10.1016/j.catena.2018.11.035

Embrapa – Empresa Brasileira de Pesquisa Agropecuária. (1997). Centro Nacional de Pesquisa de Solos. Manual de métodos de análise de solo. 2nd ed. Ministério da Agricultura e do Abastecimento, Brasília, p 212. (In Portuguese).

Freitas, L. P., Lopes, M. L., Carvalho, L. B., Panosso, A. R., La Scala Júnior, N., Freitas, R. L., & Lotufo, A. D. (2018). Forecasting the spatiotemporal variability of soil CO 2 emissions in sugarcane areas in southeastern Brazil using artificial neural networks. *Environmental monitoring and assessment*, *190*, 1–14. https://doi.org/10.1007/s10661-018-7118-0

Ghadernejad, K., Shahgholi, G., Mardani, A., & Chiyaneh, H. G. (2018). Prediction effect of farmyard manure, multiple passes and moisture content on clay soil compaction using adaptive neuro-fuzzy inference system. *Journal of Terramechanics, 77*, 49–57. https://doi.org/10.1016/j.jterra.2018.03.002

Gliński, J., & Stępniewski, W. (1985). *Soil aeration and its role for plants*. Boca Raton: CRC Press.

Golchin, A., Oades, J. M., Skjemstad, J. O., & Clarke, P. (1994). Soil structure and carbon cycling. *Soil Research, 32*(5), 1043–1068. https://doi.org/10.1071/SR9941043

Hamedi, S., & Jahromi, H. D. (2021). Performance analysis of all-optical logical gate using artificial neural network. *Expert Systems with Applications*, *178*, 115029 https://doi.org/10.1016/j.eswa.2021.115029

Hanson, P. J., Edwards, N. T., Garten, C. T., & Andrews, J. A. (2000). Separating root and soil microbial contributions to soil respiration: A review of methods and observations. *Biogeochemistry, 48*, 115–146. https://doi.org/10.1023/A:1006244819642

Haykin, S. (2001). Redes neurais: Princípios e prática. Bookman Editora. https://doi.org/10.1002/0471221546

Haykin, S. (2009). Neural networks and learning machines, 3/E. Pearson Education India.

Heddam, S. (2014). Modeling hourly dissolved oxygen concentration (DO) using two different adaptive neuro-fuzzy inference systems (ANFIS): *A comparative study. Environmental Monitoring and Assessment*, *186*(1), 597–619. https://doi.org/10.1007/s10661-013-3402-1

Huang, N., Niu, Z., Zhan, Y., Xu, S., Tappert, M. C., Wu, C., & Cai, D. (2012). Relationships between soil respiration and photosynthesis-related spectral vegetation indices in two cropland ecosystems. *Agricultural and Forest Meteorology*, *160*, 80–89. https://doi.org/10.1016/j.agrformet.2012.03.005

Hunter, A., Kennedy, L., Henry, J., & Ferguson, I. (2000). Application of neural networks and sensitivity analysis to improved prediction of trauma survival. *Computer Methods and Programs in Biomedicine, 62*(1), 11–19. https://doi.org/10.1016/S0169-2607(99)00046-2

Jang, J. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions Systems, Man, and Cybernetics*, *23*(3):665–685. https://doi.org/10.1109/21.256541

Kaab, A., Sharifi, M., Mobli, H., Nabavi-Pelesaraei, A., & Chau, K. W. (2019). Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Science of the Total Environment, 664*, 1005–1019. https://doi.org/10.1016/j.scitotenv.2019.02.004

Kharb, R. K., Shimi, S. L., Chatterji, S., & Ansari, M. F. (2014). Modeling of solar PV module and maximum power point tracking using ANFIS. *Renewable and Sustainable Energy Reviews, 33*, 602–612. https://doi.org/10.1016/j.rser.2014.02.014

Kim, D. G., Vargas, R., Bond-Lamberty, B., & Turetsky, M. R. (2012). Effects of soil rewetting and thawing on soil gas fluxes: A review of current literature and suggestions for future research. *Biogeosciences*, *9*(7), 2459-2483. https://doi.org/10.5194/bg-9-2459-2012,2012

Klir, G., & Yuan, B. (1995). Fuzzy sets and fuzzy logic (Vol. 4, pp. 1–12). New Jersey: Prentice hall.

Kooch, Y., & Ghaderi, E. (2021). Soil function can sensitively respond to different canopy composition of Crataegus and Berberis. *Applied Soil Ecology*, *167*, 104112. https://doi.org/10.1016/j.apsoil.2021.104112

Kooch, Y., & Noghre, N. (2020). The effect of shrubland and grassland vegetation types on soil fauna and flora activities in a mountainous semi-arid landscape of Iran. *Science of the total environment*, *703*, 135497. https://doi.org/10.1016/j.scitotenv.2019.135497

Kuzyakov, Y., & Gavrichkova, O. (2010). Time lag between photosynthesis and carbon dioxide efflux from soil: A review of mechanisms and controls. *Global Change Biology, 16*(12), 3386–3406. https://doi.org/10.1111/j.1365-2486.2010.02179.x

La Scala Jr, N., Panosso, A. R., & Pereira, G. T. (2003). Modelling short-term temporal changes of bare soil CO2 emissions in a tropical agrosystem by using meteorological data. *Applied Soil Ecology, 24*(1), 113–116. https://doi.org/10.1016/S0929-1393(03)00065-9

Laganière, J., Paré, D., Bergeron, Y., & Chen, H. Y. (2012). The effect of boreal forest composition on soil respiration is mediated through variations in soil temperature and C quality. *Soil Biology and Biochemistry*, *53*, 18–27. https://doi.org/10.1016/j.soilbio.2012.04.024

Lal, R. (2009). Challenges and opportunities in soil organic matter research. *European Journal of Soil Science, 60*(2), 158–169. https://doi.org/10.1111/j.1365-2389.2008.01114.x

Lamzouri, K., Latrach, L., Mahi, M., Ouattar, S., Bartali, H. E., Masunaga, T., & Mandi, L. (2017). Controlling biochemical oxygen demand in the multi-soil-layering using neural network tool. *Journal of Materials Environmental Science*. http://www.jmaterenvironsci.com/Document/vol8/vol8_N6/214-JMES-3059-Lamzouri.pdf

Li, D., Niu, S., & Luo, Y. (2012). Global patterns of the dynamics of soil carbon and nitrogen stocks following afforestation: A meta-analysis. *New Phytologist, 195*(1), 172–181. https://doi.org/10.1111/j.1469-8137.2012.04150.x

Li, M., Liu, T., Duan, L., Ma, L., Wang, Y., Zhou, Y., & Lei, H. (2021). Hydrologic gradient changes of soil respiration in typical steppes of Eurasia. *Science of The Total Environment*, *794*, 148684. https://doi.org/10.1016/j.scitotenv.2021.148684

Liang, L., Peng, S., Sun, J., Chen, L., & Cao, Y. (2010). Estimation of annual potential evapotranspiration at regional scale based on the effect of moisture on soil respiration. *Ecological Modelling, 221*(22), 2668–2674. https://doi.org/10.1016/j.ecolmodel.2010.08.010

Liptzin, D., Silver, W. L., & Detto, M. (2011). Temporal dynamics in soil oxygen and greenhouse gases in two humid tropical forests. *Ecosystems, 14*, 171–182. https://doi.org/10.1007/s10021-010-9402-x

Liu, J., Sui, Y., Yu, Z., Shi, Y. U., Chu, H., Jin, J., & Wang, G. (2014). High throughput sequencing analysis of biogeographical distribution of bacterial communities in the black soils of Northeast China. *Soil Biology and Biochemistry*, *70*, 113–122. https://doi.org/10.1016/j.soilbio.2013.12.014

Liu, X., Ji, L., Zhang, C., & Liu, Y. (2022). A method for reconstructing NDVI time-series based on envelope detection and the Savitzky-Golay filter. *International Journal of Digital Earth, 15*(1), 553–584. https://doi.org/10.1080/17538947.2022.2044397

Liu, Y., Zhang, Q., Song, L., & Chen, Y. (2019). Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. *Computers and Electronics in Agriculture*, *165*, 104964. https://doi.org/10.1016/j.compag.2019.104964

Lucas-Borja, M. E., de Santiago, J. H., Yang, Y., Shen, Y., & Candel-Pérez, D. (2019). Nutrient, metal contents and microbiological properties of litter and soil along a tree age gradient in Mediterranean forest ecosystems. *Science of the Total Environment, 650*, 749–758. https://doi.org/10.1016/j.scitotenv.2018.09.079

Mäkiranta, P., Minkkinen, K., Hytönen, J., & Laine, J. (2008). Factors causing temporal and spatial variation in heterotrophic and rhizospheric components of soil respiration in afforested organic soil croplands in Finland. *Soil Biology and Biochemistry, 40*(7), 1592–1600. https://doi.org/10.1016/j.soilbio.2008.01.009

Manly, B. F., & Alberto, J. A. N. (2008). Métodos estatísticos multivariados: uma introdução. Bookman Editora.

Moreira, F. D. S., & Siqueira, J. O. (2006). Microbiologia e bioquımica do solo. Lavras, MG, UFLA.

Neira, J., Ortiz, M., Morales, L., & Acevedo, E. (2015). Oxygen diffusion in soils: Understanding the factors and processes needed for modeling. *Chilean Journal of Agricultural Research, 75*, 35–44. https://doi.org/10.4067/S0718-58392015000300005

Nourbakhsh, F. (2007). Decoupling of soil biological properties by deforestation. *Agriculture, Ecosystems & Environment, 121*(4), 435–438. https://doi.org/10.1016/j.agee.2006.11.010

Ouyang, Y., & Zheng, C. (2000). Surficial processes and CO2 flux in soil ecosystem. *Journal of Hydrology, 234*(1–2), 54–70. https://doi.org/10.1016/S0022-1694(00)00240-7

Ozturk, M., Salman, O., & Koc, M. (2011). Artificial neural network model for estimating the soil temperature. *Canadian Journal of Soil Science, 91*(4), 551–562. https://doi.org/10.4141/cjss1007

Padarian, J., Minasny, B., & McBratney, A. B. (2019). Machine learning and soil sciences: A review aided by machine learning tools. https://doi.org/10.5194/soil-6-35-2020

Panosso, A. R., Perillo, L. I., Ferraudo, A. S., Pereira, G. T., Miranda, J. G. V., & La Scala Jr, N. (2012). Fractal dimension and anisotropy of soil CO2 emission in a mechanically harvested sugarcane production area. *Soil and Tillage Research, 124*, 8–16. https://doi.org/10.1016/j.still.2012.04.0

Paul, K. I., Polglase, P. J., Smethurst, P. J., O'Connell, A. M., Carlyle, C. J., & Khanna, P. K. (2004). Soil temperature under forests: A simple model for predicting soil temperature under a range of forest types. *Agricultural and Forest Meteorology, 121*(3–4), 167–182. https://doi.org/10.1016/j.agrformet.2003.08.030

Prado, R. B., Fidalgo, E. C. C., Monteiro, J. M. G., Schuler, A. E., Vezzani, F. M., Garcia, J. R., & Simões, M. (2016). Current overview and potential applications of the soil ecosystem services approach in Brazil. *Pesquisa Agropecuária Brasileira*, *51*, 1021–1038. https://doi.org/10.1590/S0100-204X2016000900002

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved March 5, 2022, from https://www.R-project.org/

Raich, J. W., & Tufekciogul, A. (2000). Vegetation and soil respiration: Correlations and controls. *Biogeochemistry, 48*, 71–90. https://doi.org/10.1023/A:1006112000616

Raij, B. V., Andrade, J. C., Cantarella, H., Quaggio, J. A. (2001). Análise química para avaliação da fertilidade de solos tropicais. Instituto Agronômico, Campinas, p 285. (In Portuguese).

Reichardt, K., & Timm, L. (2012). Solo, planta e atmosfera: Conceitos, processos e aplicações–Balanço hídrico. *Barueri: Manole, cap, 15*, 317–336.

Riza, L. S., Bergmeir, C., Herrera, F., & Benitez, J. M. (2015). Fuzzy rule-based systems for classification and regression in R. *Journal of Statistical Software*, *65*(6). http://www.jstatsoft.org/v65/i06/

Rolim, G. D. S., Camargo, M. B. P. D., Lania, D. G., & Moraes, J. F. L. D. (2007). Climatic classification of Köppen and Thornthwaite sistems and their applicability in the determination of agroclimatic zonning for the state of São Paulo, Brazil. *Bragantia, 66*, 711–720. https://doi.org/10.1590/S0006-87052007000400022

Rubio, V. E., & Detto, M. (2017). Spatiotemporal variability of soil respiration in a seasonal tropical forest. *Ecology and Evolution, 7*(17), 7104–7116. https://doi.org/10.1002/ece3.3267

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627–1639.

Schaap, M. G., & Leij, F. J. (1998). Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil and Tillage Research, 47*(1–2), 37–42. https://doi.org/10.1016/S0167-1987(98)00070-1

Schwendenmann, L., Veldkamp, E., Brenes, T., & O'brien, J. J., & Mackensen, J. (2003). Spatial and temporal variation in

soil CO 2 efflux in an old-growth neotropical rain forest, La Selva, Costa Rica. *Biogeochemistry, 64*, 111–128. https://doi.org/10.1023/A:1024941614919

Silva, J. R., Silva, D. J., Gava, C. A. T., Oliveira, T. C. T. D., & Freitas, M. D. S. C. D. (2016). Carbon in humic fractions of organic matter in soil treated with organic composts under mango cultivation. *Revista Brasileira de Ciência do Solo, 40*. https://doi.org/10.1590/18069657rbcs20150095

Silva-Olaya, A. M., Cerri, C. E. P., La Scala Jr, N., Dias, C. T. S., & Cerri, C. C. (2013). Carbon dioxide emissions under different soil tillage systems in mechanically harvested sugarcane. *Environmental Research Letters*, *8*(1), 015014. https://doi.org/10.1088/1748-9326/8/1/015014

Sparling, G. P. (1997). Soil microbial biomass, activity and nutrient cycling as indicators of soil health. *Biological Indicators of Soil Health*, 97–119.

Strassburg, B. B., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., & Balmford, A. (2017). Moment of truth for the Cerrado hotspot. *Nature Ecology & Evolution*, *1*(4), 0099. https://doi.org/10.1038/s41559-017-0099

Tang, J., Baldocchi, D. D., & Xu, L. (2005). Tree photosynthesis modulates soil respiration on a diurnal time scale. *Global Change Biology, 11*(8), 1298–1304. https://doi.org/10.1111/j.1365-2486.2005.00978.x

Tedesco, M. J., Gianello, C., Bissani, C. A., Bohnen, H., & Volkweiss, S. J. (1995). Análise de solo, plantas e outros materiais, 2nd edn. *Universidade Federal do Rio Grande do Sul, Porto Alegre*, 147. Boletim técnico, 5.

Teixeira, D. D. B., Bicalho, E. D. S., Panosso, A. R., Perillo, L. I., Iamaguti, J. L., Pereira, G. T., & La Scala Jr, N. (2012). Uncertainties in the prediction of spatial variability of soil CO2 emissions and related properties. *Revista Brasileira De Ciência Do Solo, 36*, 1466–1475. https://doi.org/10.1590/S0100-06832012000500010

Tormena, C. A., Silva, A. D., & Libardi, P. L. (1998). Caracterização do intervalo hídrico ótimo de um Latossolo Roxo sob plantio direto. *Revista Brasileira de Ciência do solo*, *22*, 573–581. https://doi.org/10.1590/S0100-06831998000400002

Trigueiro, W. R., Nabout, J. C., & Tessarolo, G. (2020). Uncovering the spatial variability of recent deforestation drivers in the Brazilian Cerrado. *Journal of Environmental Management*, *275*, 111243. https://doi.org/10.1016/j.jenvman.2020.111243

Vicentini, M. E., Pinotti, C. R., Hirai, W. Y., de Moraes, M. L. T., Montanari, R., Filho, M. C. M. T., & Panosso, A. R. (2019). CO 2 emission and its relation to soil temperature, moisture, and O 2 absorption in the reforested areas of Cerrado biome, Central Brazil. *Plant and Soil*, *444*, 193–211. https://doi.org/10.1007/s11104-019-04262-z

Wang, W., Peng, S., & Fang, J. (2010). Root respiration and its relation to nutrient contents in soil and root and EVI among 8 ecosystems, northern China. *Plant and Soil, 333*, 391–401. https://doi.org/10.1007/s11104-010-0354-x

Willmott, C. J. (1981). On the validation of models. *Physical Geography, 2*, 184–194. https://doi.org/10.1080/02723646.1981.10642213

Wu, C., Gaumont-Guay, D., Black, T. A., Jassal, R. S., Xu, S., Chen, J. M., & Gonsamo, A. (2014). Soil respiration mapped by exclusively use of MODIS data for forest landscapes of Saskatchewan, Canada. *ISPRS Journal of Photogrammetry and Remote Sensing, 94*, 80–90. https://doi.org/10.1016/j.isprsjprs.2014.04.018

Yang, H., Cheng, Y., & Li, G. (2021). A denoising method for ship radiated noise based on Spearman variational mode decomposition, spatial-dependence recurrence sample entropy, improved wavelet threshold denoising, and Savitzky-Golay filter. *Alexandria Engineering Journal*, *60*(3), 3379–3400. https://doi.org/10.1016/j.aej.2021.01.055

Yuan, Z. Q., Jiang, X. J., Liu, G. J., Jin, H. J., Chen, J., & Wu, Q. B. (2019). Responses of soil organic carbon and nutrient stocks to human-induced grassland degradation in a Tibetan alpine meadow. *Catena*, *178*, 40–48. https://doi.org/10.1016/j.catena.2019.03.001

Zhao, Z., Peng, C., Yang, Q., Meng, F. R., Song, X., Chen, S., Epule, T. E., Li, P., & Zhu, Q. (2017). Model prediction of biome-specific global soil respiration from 1960 to 2012. *Earth's Futur., 2017*(5), 715–729. https://doi.org/10.1002/2016EF000480

Zhou, W. J., Sha, L. Q., Schaefer, D. A., Zhang, Y. P., Song, Q. H., Tan, Z. H., Deng, Y., Deng, X. B., Guan, H. L. (2015). Direct effects of litter decomposition on soil dissolved organic carbon and nitrogen in a tropical rainforest. *Soil Biology and Biochemistry*, *81*, 255–8. https://doi.org/10.1016/j.soilbio.2014.11.019

Zounemat-Kermani, M., & Scholz, M. (2014). Modeling of dissolved oxygen applying stepwise regression and a template-based fuzzy logic system. *Journal of Environmental Engineering, 140*, 69–76. https://doi.org/10.1061/(asce)ee.1943-7870.0000780