Universidade Federal Rural do Semi-Árido
Pró-Reitoria de Pesquisa e Pós-Graduação
https://periodicos.ufersa.edu.br/index.php/caatinga

ISSN 0100-316X (impresso)
ISSN 1983-2125 (online)
http://dx.doi.org/10.1590/1983-21252022v35n424rc

ESTIMATING CO_2 EMISSIONS FROM TILLED SOILS THROUGH ARTIFICIAL NEURAL NETWORKS AND MULTIPLE LINEAR REGRESSION[1]

EDNEY LEANDRO DA VITÓRIA[2]*, CARLA DA PENHA SIMON[3], ELCIO DAS GRAÇA LACERDA[4], ISMAEL LOURENÇO DE JESUS FREITAS[5], IVONEY GONTIJO[2]

ABSTRACT - Quantifying soil gas emissions is costly, since it requires specific methodologies and equipment. The objective of this study was to evaluate modeling by nonlinear regression and artificial neural networks (ANN) to estimate CO_2 emissions caused by soil managements. CO_2 emissions were evaluated in two different soil management systems: no-tillage and minimum tillage. Readings of CO_2 flow were carried out by an automated closed system chamber; soil temperature, water content, density, and total organic carbon were also determined. The regression model and the ANN models were adjusted based on the correlation of the variables measured in the areas where the soil was managed with no-tillage and minimum tillage with data of CO_2 emission. Artificial neural networks are more accurate to determine correlations between CO_2 emissions and soil temperature, water content, density, and organic carbon content than linear regression.

Keywords: Greenhouse gases. Soil management. Modeling. Artificial intelligence.

ESTIMATIVAS DE EMISSÃO DE CO_2 EM SOLOS CULTIVADOS POR MEIO DE REDES NEURAIS ARTIFICIAS E MODELO LINEAR DE REGRESSÃO

RESUMO - A quantificação das emissões destes gases do solo é onerosa, uma vez que requer metodologias e equipamentos específicos. O objetivo deste foi avaliar a modelagem utilizando regressão não linear e redes neurais artificiais para estimar a emissão de CO_2 em função do manejo do solo, e de suas propriedades físicas e químicas. A emissão de CO_2 foi avaliada em dois diferentes manejos do solo, o plantio direto e o cultivo mínimo. As leituras de fluxo CO_2 foram realizadas por meio de uma câmara de sistema fechado automático, determinou-se ainda a temperatura e teor de água do solo, densidade do solo e carbono orgânico total. O modelo de regressão e os modelos de redes neurais artificiais foram ajustados a partir da correlação entre as variáveis medidas nas áreas em que o solo foi manejado com plantio direto e cultivo mínimo, com os dados de emissão de CO_2. As redes neurais artificiais são mais precisas na determinação das relações entre a emissão de CO_2 e a temperatura, teor de água no solo, densidade do solo e carbono orgânico, quando comparado com os resultados de regressão linear.

Palavras-chave: Gases de efeito estufa. Manejo do solo. Modelagem. Inteligência artificial.

*Corresponding author
[1]Received for publication in 01/29/2022; accepted in 05/18/2022.
[2]Department of Agricultural and Biological Sciences, Universidade Federal do Espírito Santo, São Mateus, ES, Brazil; edney.vitoria@ufes.br – ORCID: 0000-0002-2268-6037, ivoney.gontijo@ufes.br – ORCID: 0000-0002-4251-4689.
[3]Nuclear Energy Center in Agriculture, Universidade de São Paulo, Piracicaba, SP, Brazil; carlasimon@usp.br – ORCID: 0000-0003-4312-4612.
[4]Universidade Federal do Espírito Santo, Santa Teresa, ES, Brazil; elciogl@ifes.edu.br – ORCID: 0000-0002-9816-4657.
[5]Secretary of Education of Espirito Santo, Governo do Estado do Espírito Santo, Vitória, ES, Brazil; ismaelljf@yahoo.com.br – ORCID: 0000-0002-8138-7846.

INTRODUCTION

Greenhouse gas (GHG) emissions contribute to climate change and are the subject of studies on their effects, causes, and mitigation factors (OERTEL et al., 2016; RIAHI et al., 2017; NAWAZ et al., 2019). It is estimated that world agriculture accounts for approximately 22% of total carbon dioxide emissions, 80% of nitrous oxide emissions, and 55% of methane emissions (LIU et al., 2019).

The cultivation method, implementation of precision farming techniques, use of crops with high potential for carbon sequestration, adequate management of crops and pastures, and reforestation of agricultural areas are factors that can reduce GHG emissions from agriculture (ZHAO et al., 2018; RUTKOWSKA et al., 2018).

The adaptation of current land-use practices should be considered relevant to mitigate climate change, because CO_2 from basal soil respiration is important for carbon cycle. CO_2 emissions are related to soil properties: organic matter (YUSTE et al., 2019), soil moisture (DOWHOWER et al., 2020), soil temperature (YUSTE et al., 2019; ZOU et al., 2018), and soil microbial communities (NIKOLENKO et al., 2019). There is proven evidence that properly managed soils may become carbon sinks, contributing significantly to the reduction of CO_2 emissions to the atmosphere (FARINA et al., 2017).

The quantification of emissions of these soil gases is costly, since it requires specific methodologies and equipment. Some methodologies for their estimation by modeling CO_2 emissions and/ or carbon sequestration potential in agricultural soils have been tested at different scales (MARTÍN et al., 2016; GOMES et al., 2019).

The application of artificial neural networks (ANN) can be an alternative for modeling the correlation of CO_2 emissions with other parameters related to soil physical and chemical properties and climatic and environmental factors. This approach has been successfully applied to optimize, predict, and control complex systems (BURAGIENĖ et al., 2019; DIAO et al, 2021), such as CO_2 emissions. Artificial neural networks have been applied for modeling relationships between parameters within complex systems in different environmental areas, for example, to estimate soil temperature (FERNANDES et al., 2019), predict chemical composition and other soil properties from field observations (MELAKU et al., 2020) and soil NO emissions (DENG et al., 2020; JEREMIAH et al., 2021), and model soil respiration in forest ecosystems (LIMA et al., 2020).

The hypothesis considered in the present study is that the use of ANN presents higher efficiency and precision for the estimation of CO_2 emissions than multivariate regression modeling. Therefore, researches on ANN may be useful to estimate GHG because it is a technique that presents significant results for evaluations of complex systems. Thus, the objective of the present study was to evaluate modeling by nonlinear regression and ANN to estimate CO_2 emissions caused by soil managements.

MATERIAL AND METHODS

The study was conducted at the experimental area of the Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo, Santa Teresa campus (19°48'17"S, 40°40'34"W, and average altitude of 125 m), in the municipality of Santa Teresa, ES, Brazil. The climate of the region is Cwa, characterized as humid temperate, with dry winter and hot summer, according to the Köeppen classification (ALVARES et al., 2014). According to a local weather station, the mean annual rainfall depth is 1,161 mm, with mean annual temperature of 24.4 °C. The predominant soil of the study area presented a clayey texture and was classified as a Typic Hapludox (Latossolo Amarelo; EMBRAPA, 2018). The study was conducted from August 2016 to March 2017. Soil preparation was carried out in an area under a center pivot irrigation system that covered 12 hectares (ha).

The no-tillage system had been implemented in an area of 6 ha since 2009, with crop rotation (common bean - maize - velvet beans or sorghum) for grain production. Minimum tillage was implemented in an area where some conservation practices, such as fallow, had been carried out since 2015.

Soil CO_2 flow readings (μmol CO_2 m^{-2} s^{-1}) were carried out using an automated closed system chamber (LI-COR® Biosciences, Lincoln, USA), which uses an infrared gas analyzer model LI-8100A, with an opaque chamber model LI-8100-104C. This system was operated through sampling carbon dioxide concentrations by optical absorption spectroscopy. Installation of PVC rings (0.203 m diameter and 0,1143 m height) was necessary for the measurements because it prevents disturbances that the direct insertion of the chamber into the soil could cause, such as alteration in porous structures, which would directly affect the soil CO_2 emission, overestimating the readings.

The PVC rings were inserted into the soil with a chamber off-set (of the top of the ring above the soil surface) of 0.02 m; the dry matter on the soil was cut, so that the ring fixation does not modify the soil cover composition. The PVC rings were installed at 180 days before the readings. The readings were carried out in three consecutive days in each system. Each reading lasted 2 min, measuring the carbon dioxide concentration inside the chamber every second.

Soil temperature and water contents were sampled simultaneously with CO_2 emission readings. A temperature and humidity sensor (5TM Decagon Devices®) was used. Sampling was carried out at 0.10 m distant from the external part of the PVC ring, reaching 0.05 m depth. The soil density in the 0.00-0.20 m layer was evaluated using an Uhland sampler and a volumetric ring.

Soil samples were dried at 50 °C for 24 hours to determine total organic carbon (TOC). Roots and other plant residues were removed, and the remaining plant material was removed by flotation in 0.01 M HCl and sieved at 210 µm. TOC was determined in a Carlo Erba Analyzer (CHN-1110) coupled to a mass spectrometer (Optima Thermo Finnigan Plus Delta). The analyses were carried out at the Laboratório Agronômico de Análise de Solo, Folha e Água (LAGRO) of the Centro Universitário Norte do Espírito Santo of the Universidade Federal do Espírito Santo (UFES). Analytical uncertainties ranged, on average, from 0.3% to 0.5%. TOC contents were expressed in grams per kilogram (g kg^{-1}) of dry soil.

The regression model was adjusted for the correlation analysis between the measured variables (soil temperature, soil water content, soil density, and TOC in areas where the soil was managed with no-tillage and minimum tillage) and the CO_2 emission data. Multiple linear regression procedures, by the least square method, and ANN were used to estimate the CO_2 emissions flow for the two soil management systems.

The data were randomly divided into two sets to validate the methods and evaluate their performance: one for the adjustment of the regression model and training of networks (70%), and the other for validation of the regression and trained networks (30%). According to Ding, Wang, and Han (2019), the validation procedure should be applied to verify the ability of a neural network to produce adequate outputs for inputs that were not present during the training.

The regression model adopted was:

$$Y = \beta_0 + \beta_1 \cdot T + \beta_2 \cdot SH + \beta_3 \cdot SD + \beta_4 \cdot TOC$$

where Y = CO_2 emission flux; T = soil temperature; SH = soil water content; SD = soil density; TOC = total organic carbon; and βi = estimators of parameters to be adjusted.

Artificial neural networks are indicated for estimating non-linear mathematical functions or relations. They are a logical computational system made up of numerous simple processing layers linked together. In each one or more layers, there are several units interconnected by a large number of connections, usually unidirectional (BRAGA; CARVALHO; LUDEMIR, 2000). According to Haykin (2007), ANN are based on systems of natural functioning of the human brain. An ANN may consist of one or more layers, and each layer can contain one or more neurons (single-processing units). The input layer only receives the values (quantitative or qualitative) of the provided variables and transmits them to the intermediate layer. The intermediate, or hidden, layer and the output layer map the knowledge, processing information with their neurons, also called computing nodes. A computing node k receives the input signals (x_i) and assigns weights to them (w_{ki}); a sum is obtained by adding the inputs multiplied by their respective weights and adding a prefixed signal (b_k). The result of this sum (v_k) is subjected to an activation function [$f(v_k)$] and provides the output of the neuron (y_k).

The ANN training starts by presenting data (input and output variables) to a pre-established, or not, structure, depending on the software used. The training process starts with random weight values, and based on these values, the first output is compared with the respective actual value of the first observation. The difference between the output estimated by the network and the actual value generates an error signal that calibrates the weight adjustment, thus initiating a new cycle to approximate the output to the desired result, i.e., minimizing the error (HAYKIN, 2007). The basic mathematical model of an artificial neuron is presented in Equation 1.

$$y_k = \varphi\left(\sum x_m w_m\right) \tag{1}$$

where y_k = output of the artificial neuron, φ = activation function, x_m = number of inputs, and w_m = weight for each m input.

CO_2 emissions were estimated according to Ding, Wang, and Han (2019) by training multilayer perceptron networks (MLP) that have a universal ability to approximate functions. The numerical variables were normalized linearly in intervals of 0 to 1, and the categorical variables were subjected to a transformation called codification, i.e., each variable received a numerical code that enabled the calculation of the artificial neuron.

The software Neuro 4.0.6 was used in the training and validation of ANN. Four ANN were trained; three neurons in the input layer, three in the hidden layer, and one in the output layer. The numerical inputs used were soil temperature (T), soil water content (SH), soil density (SD), and TOC, and the categorical (non-numerical) variable consisted of the two soil management systems (no-tillage and minimum tillage).

The Resilient Propagation RPROP+ training was used, with a sigmoidal activation function in the hidden and output layers (Equation 2).

$$\varphi(v) = \frac{1}{1+\exp^{\phi u}} \quad (2)$$

where φ = sigmoidal activation function, $\emptyset$ = the parameter estimate that determines inclination of the sigmoidal function, and u = function activation potential.

This function was chosen because it is the most common used in the construction of ANN (DING; WANG; HAN 2019). The Resilient Propagation training algorithm represents a variant of the backpropagation algorithm (backpropagation error) and has the advantage of being able to calculate and acquire information about a given problem, because its weight adjustment depends more on the signal of error gradients and it is more efficient and recommended for ANN of the MLP type.

The criterion for finishing the training of networks was defined according to Leal et al. (2015), i.e., a total number of cycles of 3,000 or a mean square error of less than 1%. The training was finished when one of the criteria was reached.

The significance of the coefficients of each model tested was evaluated by the Student's t test at 5% significance level. Then, Pearson's linear correlation coefficient (r) and coefficient of determination (R^2) were calculated for the estimated and observed values. The significance of r was evaluated by the Student's t test at 5% significance level. The mean absolute error (MAE), root mean square error (RMSE), and Willmott's d index (Equations 3, 4, and 5) were also calculated for each model.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{c}_i - c_i)^2}{n}} \quad (3)$$

$$d = \left[\frac{\sum_{i=1}^{n}(\hat{c}_i - c_i)^2}{\sum_{i=1}^{n}\left(|\hat{c}_i - \bar{c}| + |c_i - \bar{c}|\right)^2} \right] \quad (4)$$

$$rc\hat{c} = \frac{\operatorname{cov}(c_i, \hat{c}_i)}{\sqrt{S^2(c_i)S^2(\hat{c}_i)}} \quad (5)$$

where $\hat{c}_i$ are the estimated CO_2 emission values, c_i are the observed CO_2 emission values, and $\bar{c}$ is the mean of the observed values. Subsequently, the Camargo and Sentelhas (CS) index of the product between r and d (CS = r × d) was obtained.

The criterion used to indicate the model that best estimates CO_2 emissions was the Pearson's linear correlation coefficient ($r_{c\hat{c}}$), with RMSE closer to 0, and d and CS indexes closer to 1.

The network configurations and analyses were carried out using the specific neural network tool of the software IBM-SPSS 22. The dataset was pre-processed in an Excel spreadsheet.

RESULTS AND DISCUSSION

The mean, minimum, maximum, standard deviation, and results of the normality test of CO_2 emission flux, soil temperature, soil water content, soil density, and TOC in each soil management system are presented in Table 1. The descriptive analysis allowed to explore and analyze the values of variables and verify the existence of outliers or discrepant values, as well as their influence on data of position and dispersion measurements; therefore, when identified, discrepant values were eliminated from the analysis.

None of the variables presented normal distribution by the Shapiro-Wilk test at 5% significance. In the no-tillage system, soil temperature, soil density, and TOC presented low variability in relation to the mean; CO_2 emissions and soil water content presented intermediate variability, according to the classification: low for CV <12%, intermediate for 12% < CV < 60%, and high for CV> 60%. In the minimum-tillage system, soil temperature and density presented low variability in relation to the mean, and other variables presented intermediate variability.

Table 2 shows the correlation level between the variables analyzed for the no-tillage and minimum-tillage systems. The use of soil temperature and TOC as predictor variables is justified by their significant correlations with CO_2 emission. The increase in these variables results in an increase in CO_2 emission. TOC showed stronger correlation than soil temperature in the no-tillage (r = 0.79, p<0.05) and minimum-tillage (r = 0.66, p<0.05) systems, whereas the correlation between temperature and CO_2 emissions was 0.58 (p<0.05) and 0.47 (p 0.05) for the no-tillage and minimum tillage systems, respectively. Soil water content also showed significant correlation in no-tillage (r = −0.58, p<0.05) and minimum tillage (r = −0.47, p<0.05).

Table 1. Descriptive statistics of soil CO_2 emissions (C-CO_2), soil temperature (T), soil water content (SH), soil density (SD), and total organic carbon (TOC) (g kg^{-1}) areas under no-tillage (NTS) and minimum tillage (MTS) management systems.

Management system	Variable	Minimum	Mean	Maximum	CV (%)	W
NTS	C-CO_2 (μmol m^{-2})	0.88	2.30a	5.98	37.0	0.043*
	T (°C)	27.6	32.4a	38.0	10.3	0.040*
	SH (m^3 m^{-3})	0.040	0.107a	0.163	12.1	0.032*
	SD (g cm^{-3})	1.22	1.37a	1.48	7.71	0.015*
	TOC (g kg^{-1})	8.23	15.30a	24.2	10.6	0.022*
MTS	C-CO_2 (μmol m^{-2})	0.92	2.25a	6.01	40.2	0.049*
	T (°C)	26.9	35.7a	38.8	11.2	0.051*
	SH (m^3 m^{-3})	0.040	0.093a	0.156	15.0	0.025*
	SD (g cm^{-3})	1.26	1.43a	1.47	7.69	0.019*
	TOC (g kg^{-1})	8.25	12.3b	26.1	12.7	0.005*

CV - coefficient of variation; W - Shapiro-Wilk normality test.
Means followed by the same letter do not differ by the Tukey's test ($p<0.05$) comparing the same variable between the different soil managements.
*non-normal distribution.

Table 2. Pearson's correlation coefficients for soil CO_2 emissions (C-CO_2), soil temperature (T), soil water content (SH), soil density (SD), and total organic carbon (TOC; g kg^{-1}) in areas under no-tillage and minimum-tillage management systems.

	C-CO_2	T	SH	SD	TOC
No-tillage system					
C-CO_2	1				
T	0.58*	1			
SH	−0.25*	0.58*	1		
SD	0.15^{ns}	−0.25*	−0.40*	1	
TOC	0.79*	0.15^{ns}	0.11^{ns}	-0.30^{ns}	1
Minimum-tillage system					
C-CO_2	1				
T	0.63*	1			
SH	−0.31*	0.47*	1		
SD	0.19^{ns}	-0.25^{ns}	−0.36*	1	
TOC	0.66*	0.18^{ns}	0.18^{ns}	-0.22^{ns}	1

*Significant by the t test ($p<0.05$).

Unlike TOC and temperature, increases in soil water content decreases CO_2 emission, thus conferring negative values for the correlation. Soil temperature, soil water content, and TOC are simple to obtain, with a relatively low cost. Total organic carbon is the most important variable to be correlated to soil CO_2 emission, which is positively affected by increases in TOC. The soil biological activity is connected to TOC contents (ZHANG et al., 2019). Soil management systems that maintain soil cover have positive correlation between annual soil respiration and organic carbon; therefore, CO_2 emissions are affected (ZHANG et al., 2019, VAZQUEZ et al. 2019, NIKOLENKO et al., 2019).

The maximum CO_2 emission was observed in the minimum-tillage (6.01 μmol m^{-2}), and the minimum was observed in the no-tillage system; however, there was no significant difference between the means of CO_2 emissions in the different management systems. In the experimental period, the region where the experiment was carried out underwent a severe water crisis, which contributed to the occurrence of significant differences in CO_2 emissions and in the other variables analyzed. Different studies have reported controversial data on soil CO_2 emission. Chaplot et al. (2012) and Buragienė et al. (2019) found lower CO_2 emissions for no-tilled and minimally tilled soils when compared to tilled soil areas. Huang et al. (2018), Xavier et al. (2019), Wang et al. (2020), and Shakoor et al. (2021) found higher CO_2 emissions for no-tillage systems.

The type of soil preparation for planting affects soil characteristics, depending on the different soil preparation technologies used and experimental year, in the same way as before the soil preparation. Therefore, soils managed in the same way for subsequent years present different temperature variation, TOC, and CO_2 emissions when compared to recent soil preparation systems. Significant differences were found between soil temperatures and the results presented this trend, corroborating the results found by Buragienė et al. (2019).

The coefficients of the adjusted equations found for the two soil management systems are shown in Table 3. The coefficients β1 and β4 found for soil temperature and TOC were significant for CO_2 emission in both no-tillage and minimum-tillage systems; the high β4 indicated a higher effect of organic carbon than soil temperature. The low soil mobilization in these managements protects the soil from degradation. In addition, the maintenance of crop residues on the soil surface increases the soil organic matter content (SILVA et al., 2016; SILVA et al., 2019).

The β2 and β3 coefficients found for soil water content and soil density were not significant. The correction coefficient between observed and estimated values was 0.706 for the dataset used in the adjustment, and 0.691 for the dataset used in the validation. The estimates were statistically equal to the values observed by the t test; thus, the model used is efficient for estimating the dependent variable (CO_2 emission) as a function of the independent variables (soil temperature and TOC).

Table 3. Adjustment parameters and statistics of the multiple linear regression model used in the estimation of CO_2 emission.

Management system	β_0	β_1	β_2	β_3	β_4	Adjustment		Validation		t test
						$r\,c\hat{c}$	RMSE (%)	$r\,c\hat{c}$	RMSE $_{\%}$	
NTS	5.325*	0.021*	-2.832^{ns}	0.066^{ns}	0.165*	0.706	18.193	0.691	18.424	0.232^{ns}
MTS	6.023*	0.033*	-3.118^{ns}	0.102^{ns}	0.213*					

NTS = no tillage system; MTS = minimum tillage system; β_0, β_1, β_2, β_3, and β_4 = regression coefficients; $r\,c\hat{c}$ = correlation between observed and estimated CO_2 emissions; RMSE = root mean square error in percentage; NS = not significant ($p<0.05$).

The residual dispersion graphs of the multiple linear regression model showed that the dynamic of the equation was similar for both adjustment and validation, with values properly distributed; however, an overestimation trend was observed for high CO_2 values. Most error frequencies were between −20% and 20% error, which can be considered a reasonable and acceptable distribution (Figure 1).

The four trained networks presented similar statistical precisions for both correlation and RMSE% in the training and validation (Table 4).

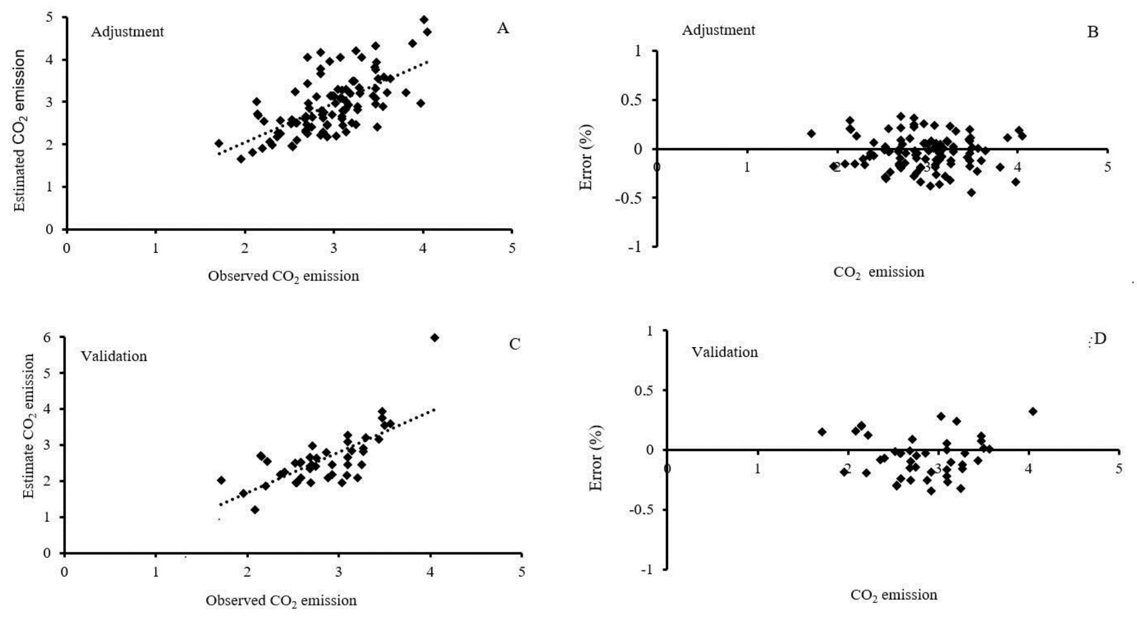

Figure 1. Observed versus estimated CO_2 emissions and residue for estimates obtained through the multiple regression model.

Table 4. Statistical precisions of artificial neural networks (ANN) trained for estimating CO_2 emission.

ANN	Training		Validation		t test
	r $c\hat{c}$	RMSE%	r $c\hat{c}$	RMSE%	
1	0.863	9.12	0.849	9.47	0.238^{ns}
2	0.849	9.17	0.848	9.51	0.253^{ns}
3	0.851	9.23	0.849	9.49	0.227^{ns}
4	0.862	9.34	0.847	9.59	0.292^{ns}

r $c\hat{c}$ = correlation between observed and estimated CO_2 emissions; RMSE% = root mean square error (%); NS = not significant ($p<0.05$).

The observed and estimated values were statistically equal by the t test, showing the accuracy of the estimates found through the four ANN. However, those of ANN 1 were slightly higher in the training when compared to the others, presenting the highest correlation coefficient (0.863) and lowest RMSE (9.12%), as well as in the validation, also presenting the lowest RMSE (9.47%). Thus, the graphs of observed versus estimated CO_2 emissions and residues were evaluated for ANN 1 (Figure 2).

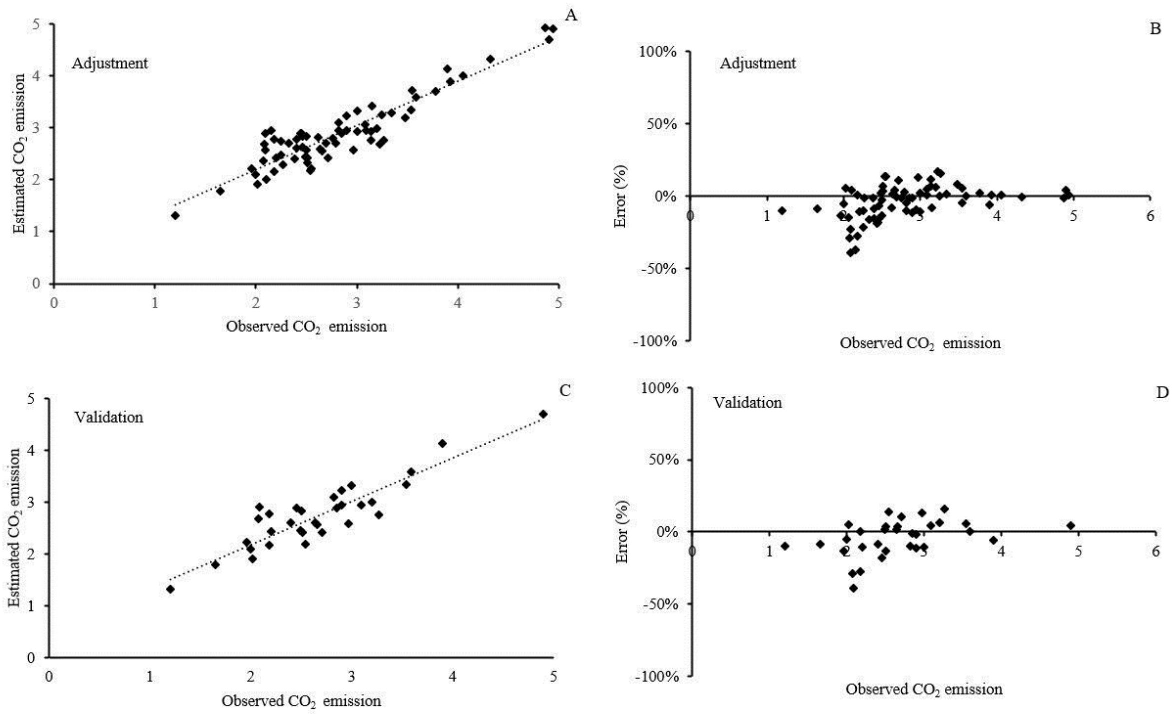

Figure 2. Observed versus estimated CO_2 emissions and residue for the estimates obtained through ANN 1.

ANN 1 presented the best training and validation statistics and similar dynamics to the estimates obtained through regression, considering the observed versus estimated values and distribution of residuals as a percentage of errors for the adjustment and validation. Thus, both techniques are efficient for estimating CO_2 emission. However, the Pearson's correlation coefficient and RMSE obtained through the network for both training and validation were higher than the multiple linear regression, indicating that ANN offer higher accuracy. Table 5 presents a summary of the statistical indicators of the two modeling techniques used.

Although the two techniques presented discrepant statistical indicators, the Camargo and Sentelhas indexes indicated that the ANN modeling is more accurate than the multiple linear regression model; the closer to 1, the better the representativeness of these indexes when compared to the other models. Unlike the multiple linear regression model, which was adjusted as a function of each soil management, there was no data stratification for the network training, which is a great differential of ANN. The possibility of inserting categorical (non-numerical) variables in the adjustment generated precise and unbiased results.

Table 5. Adjustment and precision statistics obtained through the multiple linear regression model (MLR) and artificial neural network 1 (ANN1).

	Adjustment				Validation			
	$r_{c\hat{c}}$	RMSE%	d	CS	$r_{c\hat{c}}$	RMSE%	d	CS
MLR	0.706	18.19	0.711	0.501	0.691	18.42	0.694	0.479
ANN 1	0.863	9.12	0.862	0.733	0.849	9.47	0.851	0.722

$r_{c\hat{c}}$ = correlation between observed and estimated CO_2 emissions; RMSE% = root mean square error (%); d = Willmott's index; CS = Camargo and Sentelhas index.

Vedaraman et al. (2017), Rubio and Detto (2017), and Thangavel et al. (2018) evaluated soil CO_2 emissions through neural networks under different conditions from those of the present work; the results were similar regarding the better efficiency of the ANN model, compared to other mathematical or statistical modeling processes. The higher accuracy of neural networks, when compared to linear models, is related to several factors, such as the dimensioning and precision of data collection (BURAGIENĖ et al., 2019; DIAO et al, 2021; BELCAVELLO et al., 2022), number of predictor variables, possibility of merging quantitative and qualitative variables in the same model (SARKAR; MISHRA, 2018), and possibility of a better training of neural networks due to the number of neurons used and number of hidden layers (FERNANDES et al., 2019; DENG et al., 2020; JEREMIAH et al., 2021; LACERDA et al., 2022).

CONCLUSIONS

Artificial neural networks are more accurate for determining correlations between CO_2 emissions and soil temperature, soil water content, soil density, and soil organic carbon content when compared to results of mixed linear regression. The approach of soil CO_2 emissions by artificial neural network simulation techniques is a useful and effective tool. CO_2 flux from the soil to the atmosphere can be modeled with high accuracy; deep artificial neural networks may have higher efficiency in similar works. These results show a good potential of this methodology to be applied in this type of problem.

ACKNOWLEDGMENT

This study was partly financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES 18/2020) and the Fundação de Amparo a Pesquisa do Espírito Santo (FAPES 140/2021) – Brazil. CAPES/FAPES Cooperation - Postgraduate Development Program – PDPG.

REFERENCES

ALVARES, C. A. et al. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, 22: 711-728, 2014.

BRAGA, A. P.; CARVALHO A. P. L. F.; LUDEMIR, T. B. **Redes Neurais Artificiais**: Teoria e Aplicações. Belo Horizonte, MG: LTC - Livros Técnicos e Científicos Editora S.A., 2000. 262 p.

BELCAVELLO, M. O. et al. Artificial neural networks in estimating the productivity of a forest harvesting machine. **Journal of Engineering**, 12: 52 -58, 2022.

BURAGIENĖ, S. et al. Relationship between CO_2 emissions and soil properties of differently tilled soils. **Science of the Total Environment**, 662: 786-795, 2019.

CHAPLOT, V. et al. Water erosion-induced CO2 emissions from tilled and no-tilled soils and sediments. **Agriculture, Ecosystems & Environment**, 159: 62-69, 2012.

DENG, L. et al. Soil GHG fluxes are altered by N deposition: New data indicate lower N stimulation of the N2O flux and greater stimulation of the calculated C pools. **Global Change Biology**, 26: 2613-2629, 2020.

DIAO, W. et al. Influences of soil bulk density and texture on estimation of surface soil moisture using spectral feature parameters and an artificial neural network algorithm. **Agriculture**, 11: 1- 20, 2021.

DING, D.; WANG, Z.; HAN, Q. L. Neural-network-based output-feedback control with stochastic communication protocols. **Automatica**, 106: 221-229, 2019.

DOWHOWER, S. L. et al. Soil greenhouse gas emissions as impacted by soil moisture and temperature under continuous and holistic planned grazing in native tallgrass prairie. **Agriculture**,

Ecosystems & Environment, 287: 1-11, 2020.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Sistema brasileiro de classificação de solos**. 5. ed. Brasília, DF: Embrapa Solos, 2018. 353 p.

FARINA, R. et al. Modeling regional soil C stocks and CO_2 emissions under Mediterranean cropping systems and soil types. **Agriculture, Ecosystems and Environment**, 238: 128-141, 2017.

FERNANDES, M. M. H. et al. Estimation of soil organic matter content by modeling with artificial neural networks. **Geoderma**, 350: 46-51, 2019.

GOMES, L. C. et al. Modelagem e mapeamento dos estoques de carbono orgânico do solo no Brasil. **Geoderma**, 340: 337-350, 2019.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre, RS, 2007. 900 p.

HUANG, Y. et al. Greenhouse gas emissions and crop yield in no-tillage systems: A meta-analysis. **Agriculture, Ecosystems & Environment**, 268: 144-153, 2018.

JEREMIAH, J. J. et al. Results of Application of Artificial Neural Networks in Predicting Geo-Mechanical Properties of Stabilised Clays - A Review. **Geotechnics**, 1: 147-171, 2021.

LACERDA, L. C. et al. Prediction of Mechanical Availability in Mechanized Eucalyptus Forest Harvesting Using Artificial Neural Networks. **Journal of Agricultural Science**, 14: 1-7, 2022.

LEAL, F. A. et al. Redes neurais artificiais na estimativa de volume em um plantio de eucalipto em função de fotografias hemisféricas e número de árvores. **Revista Brasileira de Biometria**, 33: 233-249, 2015.

LIMA, J. R. S. et al. Impacts of land-use changes on soil respiration in the semi-arid region of Brazil. **Revista Brasileira de Ciência do Solo**, 44: e0200092, 2020.

LIU, Y. et al. Emission mechanism and reduction countermeasures of agricultural greenhouse gases–a review. **Greenhouse Gases: Science and Technology**, 9: 160-174, 2019.

MARTÍN, J. R. et al. Assessment of the soil organic carbon stock in Spain. **Geoderma**, 264: 117-125, 2016.

MELAKU, N. D. et al. Predicting nitrous oxide emissions after the application of solid manure to grassland in the United Kingdom. **Journal of Environmental Quality**, 49: 1-13, 2020.

NAWAZ, A. et al. Rice–wheat cropping systems in South Asia: issues, options and opportunities. **Crop and Pasture Science**, 70: 395-427, 2019.

NIKOLENKO, O. et al. Dynamics of greenhouse gases in groundwater: hydrogeological and hydrogeochemical controls. **Applied Geochemistry**, 105: 31-44, 2019.

OERTEL, C. et al. Greenhouse gas emissions from soils - A review. **Geochemistry**, 76: 327-352, 2016.

RIAHI, K. et al. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. **Global Environmental Change**, 42: 153-168, 2017.

RUTKOWSKA, B. et al. Impact of reduced tillage on CO2 emission from soil under maize cultivation. **Soil and Tillage Research**, 180: 21-28, 2018.

RUBIO, V. E.; DETTO, M. Spatiotemporal variability of soil respiration in a seasonal tropical forest. **Ecology and Evolution**, 7: 7104-7116, 2017.

SARKAR, T.; MISHRA, M. Soil erosion susceptibility mapping with the application of logistic regression and artificial neural network. **Journal of Geovisualization and Spatial Analysis**, 2: 1-17, 2018.

SHAKOOR, A. et al. A global meta-analysis of greenhouse gases emission and crop yield under no-tillage as compared to conventional tillage. **Science of The Total Environment**, 750: 1-16, 2021.

SILVA, B. O. et al. Soil CO_2 emission and short-term soil pore class distribution after tillage operations. **Soil and Tillage Research**, 186: 224-232, 2019.

SILVA, F. R. D. et al. Physical properties of a Hapludox after three decades under different soil management systems. **Revista Brasileira de Ciência do Solo**, 40: e0140331, 2016.

THANGAVEL, R. et al. Evaluating organic carbon fractions, temperature sensitivity and artificial neural network modeling of CO_2 efflux in soils: Impact of land use change in subtropical India (Meghalaya). **Ecological Indicators**, 93: 129-141, 2018.

VAZQUEZ, E. et al. Effects of no-tillage and liming amendment combination on soil carbon and nitrogen mineralization. **European Journal of Soil Biology**, 93: 103090, 2019.

VEDARAMAN, N. et al. Ultrasonic extraction of natural dye from Rubia Cordifolia, optimisation using response surface methodology (RSM) & comparison with artificial neural network (ANN) model and its dyeing properties on different substrates. **Chemical Engineering and Processing: Process Intensification**, 114: 46-54, 2017.

WANG, H. et al. No tillage increases soil organic carbon storage and decreases carbon dioxide emission in the crop residue-returned farming system. **Journal of Environmental Management**, 261: 1-7, 2020.

XAVIER, C. V. et al. Crop rotation and succession in a no-tillage system: implications for CO2 emission and soil attributes. **Journal of Environmental Management**, 245: 8-15, 2019.

YUSTE, J. C. et al. Cascading effects associated with climate-change-induced conifer mortality in mountain temperate forests result in hot-spots of soil CO2 emissions. **Soil Biology and Biochemistry**, 133: 50-59, 2019.

ZHANG, Y. et al. Spatial identification of land use functions and their tradeoffs/synergies in China: Implications for sustainable land management. **Ecological Indicators**, 107: 1-14, 2019.

ZHAO, R. et al. Impacts of water and land resources exploitation on agricultural carbon emissions: The water-land-energy-carbon nexus. **Land Use Policy**, 72: 480-492, 2018.

ZOU, J. et al. Differential responses of soil CO2 and N2O fluxes to experimental warming. **Agricultural and Forest Meteorology**, 259: 11-22, 2018.