Professional paper
10.7251/AGREN2601031F
UDC 63:551.583

HYBRID MODEL OF ARTIFICIAL NEURAL NETWORKS AND PRINCIPAL COMPONENT DECOMPOSITION FOR PREDICTING GREENHOUSE GAS EMISSIONS IN THE BRAZILIAN REGION OF MATOPIBA

Milena Monteiro FEITOSA[1], José de Jesus Sousa LEMOS[2*]

[1]Federal University of Ceará (UFC), Brazil
[2]Department of Agricultural Economics, Federal University of Ceará (UFC), Brazil
*Corresponding author: lemos@ufc.br

ABSTRACT

Greenhouse gas (GHG) emissions in agricultural production represent a global environmental challenge, requiring an understanding of the factors that influence them in order to develop sustainable practices. The general objective is to investigate the factors that influence GHG emissions and reductions in agricultural production in the MATOPIBA region, from 2006 to 2017. The methodology employed involved factor analysis with decomposition into principal components and the use of Artificial Neural Networks (ANNs) to assess interactions between variables. The data was obtained from the Agricultural Census, MapBiomas, SEEG and NOAA, considering indicators such as vegetation cover, cattle numbers, pesticide use, climate variability and industrial GDP. The results showed that 70.3% of the municipalities showed an increase in GHG emissions, with the Economic-Industrial and Agricultural Practices Effect being the main factor associated with the growth in emissions, while the Livestock Intensification Effect was the main factor associated with their reduction. It is concluded that sustainable agricultural practices, combined with efficient livestock management and the maintenance of vegetation cover, are fundamental to minimizing GHG emissions in the MATOPIBA region.

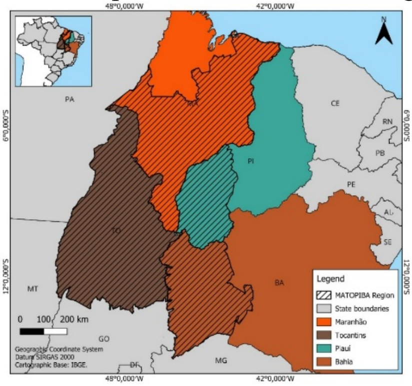
Keywords: *Greenhouse Gas Emissions (GHG); Agricultural Production; MATOPIBA; Artificial Neural Networks (ANN); Sustainability.*

INTRODUCTION

The expansion of Brazilian agriculture into new frontiers, such as the MATOPIBA region (Maranhão, Tocantins, Piauí, and Bahia) (Figure 1), has established the country as a global leader but has come at a significant environmental cost (Santos & Naval, 2022). This region, located predominantly in the Cerrado biome, has become an epicenter of land conversion, accounting for nearly half of Brazil's native vegetation deforestation in 2023 (RAD, 2023). This land-use change is directly linked to increased Greenhouse Gas (GHG) emissions (Noojipady et al., 2017;

Rausch et al., 2019), creating an urgent need to understand the factors that drive both the rise and potential reduction of these emissions (Bezerra, 2022).

Figure 1. Location Map of the MATOPIBA Region in Brazil



*Source: Developed by the author.

Based on the above, this research aims to answer the following questions: 1 - How many and which municipalities in MATOPIBA had an increase or decrease in GHG emissions between 2006 and 2017? 2 - Which variables, and in what proportions, probably influenced these emissions in this time interval?

In order to answer these questions, the general objective of this study is to investigate the factors that influence GHG emissions and reductions in agricultural production in the MATOPIBA region between 2006 and 2017. Specifically, the study seeks to: a - Ascertain the number of municipalities in the MATOPIBA region and, by state, identify which had an increase or decrease in GHG emissions between 2006 and 2017; b - Analyze the interaction between the variables tested in determining GHG emissions in this period; c - Evaluate how GHG emissions are influenced by the synergies between the indicators analyzed.

MATERIAL AND METHODS

Database and construction of indicators

The research uses secondary data extracted from the 2006 and 2017 Agricultural Censuses, MapBiomas, NOAA (National Oceanic and Atmospheric Administration) and SEEG (System of Estimates of Greenhouse Gas Emissions and Removals), from which information was obtained on the variables that are supposed to affect greenhouse gas emissions in the municipalities of the MATOPIBA region over this 11-year period. The variables and data sources used are shown in Table 1.

Table 1. Variables that, by hypothesis, affect greenhouse gas (GHG) emissions positively (+) or negatively (-) between 2006 and 2017 in MATOPIBA in this research

Variables	Hypothesis of the relation between Yi and Xij	Definition	Sources
Y_i	GHG Emissions	(GHG_{2017}/GHG_{2006}) Emissions	Greenhouse Gas Emissions - SEEG (OC, 2022).
X_{i1}	(-)	(Municipal average annual rainfall2017) / (Municipal average annual rainfall2006)	NOAA (2022)
X_{i2}	(-)	(Vegetation cover) [(Crop areas + forest areas) / (total establishment area 2017)] / [(Crop areas + forest areas) / (total establishment area 2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i3}	(-)	[(Agricultural production value2017) / (Harvested agricultural area2017)] / [(Agricultural production value2006) / (Harvested agricultural area2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i4}	(-)	[(Livestock production value2017) / (Pasture area2017)] / [(Livestock production value2006) / (Pasture area2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i5}	(-)	[(Recovered areas2017) / (Deforested areas2017)] / [(Recovered areas2006) / (Deforested areas2006)]	MapBiomas
X_{i6}	(+)	[(Cattle quantity2017) / (pasture areas 2017)] / [(Cattle quantity2006) / (pasture areas 2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i7}	(+)	[(Total tractors and machinery2017) / (total establishment area2017)] / [(Total tractors and machinery2006) / (total establishment area2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i8}	(+)	[(Expenditure on agricultural pesticides2017) / (total area of municipal establishments 2017)] / [(Expenditure on agricultural pesticides2006) / (total area of municipal establishments 2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i9}	(+)	[(Industrial sector GDP2017) / (Total municipal GDP2017)] / [(Industrial sector GDP2006) / (Total municipal GDP2006)]	Agricultural Censuses of 2006 and 2017 / IBGE
X_{i10}	(+)	CV rainfall 2017/ CV rainfall 2006	NOAA (2022)
X_{i11}	(+)	(Burn scars areas 2017 / Burn scars areas 2006)	MapBiomas

Source: Compiled based on data from SEEG, MapBiomas, Agricultural Census (2006 and 2017), IBGE, and NOAA (2022).

The methodological approaches adopted to achieve the objectives of this research begin with the construction of the used indicators. In order to assess whether there was a change in greenhouse gas (GHG) emissions between 2006 and 2017, as well as the impact of the variables shown in Table 1 on these emissions, the indicators are created as follows: the relationship between the values observed in 2017 (final year) and 2006 (initial year) is estimated for both GHG emissions and the explanatory variables.
This makes it possible to identify whether each variable increased or decreased over the period analyzed. In municipalities where the ratio between GEE_{2017} and GEE_{2006} is greater than 1, there has been an increase in emissions; if it is less than 1, there has been a reduction. The same process is applied to the explanatory or independent variables (Table 1)

Methodology adopted to assess the first research objective
In order to achieve the first objective of the research, we estimated the totals of municipalities in which the ratios between GHG emissions in 2017 were higher than those observed in 2001. In these cases, the ratios Yi_{2017}/ Yi_{2006}. The relationships between the variables that are thought to have influenced GHG emissions between 2006 and 2017 are also measured. In this case, the relationships X_{ij2017}/ X_{ij2006} are measured.

Methodology for achieving the second objective
In order to estimate the synergy between the variables that are thought to have influenced GHG emissions, Factor Analysis (FA) was used, using the principal component decomposition technique.
Before using the principal component decomposition model, it was decided to transform all the variables into indices. The indices range from 1 to 100. In the case of the dependent variable, the ratio of GHG emissions between 2006 and 2017, the following procedure was adopted. The municipalities were ranked in descending order by the ratio of GHG emissions between 2006 and 2017. Therefore, the higher the value of this ratio, the higher the GHG emissions between 2006 and 2017. For this reason, the highest emissions value was assigned the index=100. The other values were adjusted proportionally using a simple, straightforward rule of three. Thus, in the municipality where the GHG emissions index = 100, there was the highest emission of this gas between 2006 and 2017. In those municipalities where the GHG index is close to 1, this means that there was the greatest reduction in these emissions.
With regard to the 11 independent variables used to cause GHG emissions, the following criteria were adopted. All 5 variables whose hypothesis in this study establishes that they should cause a reduction in GHG emissions

between 2006 and 2017 ($GHG_{2017/2006} \leq 1$) were ranked in ascending order. To the lowest value (worst case) is assigned an index of 100. The remaining values are adjusted proportionally using a simple inverse rule of three. These variables are marked with a (-) sign in Table 1 indicating that, by hypothesis, they cause a reduction in GHG emissions.

The other 6 independent variables which, by hypothesis, should cause an increase in GHG emissions ($GHG_{2017/2006}$) were ranked in descending order. The highest value of these variables (worst case) was assigned an index of 100. The other values are adjusted proportionally using a simple, direct rule of three. These variables are marked with a (+) sign in Table 1 indicating that, by hypothesis, they cause an increase in GHG emissions.

Methodology adopted to achieve the third objective

To achieve the third objective of this research, the Artificial Neural Networks (ANN) model was used in order to investigate how GHG emissions are influenced by the synergy of partial indices (I_1, I_2, ..., I_p). ANNs are part of computational artificial intelligence. One of the main areas of application of ANNs is in the prediction of multivariate statistical data that is both non-linear and non-parametric (Sharda & Patil, 1992; Lee et al., 2017).

Zhang et al. (1998) reports that one of the procedures of computational artificial intelligence normally used to predict time series is the training of ANNs, based on the architecture and learning of the human brain. In this way, according to Zhang et al (1998), ANNs work like the human brain, seeking to recognize regularities and patterns in data, being able to learn from experience and make generalizations based on previously accumulated knowledge. ANNs are non-linear models, unlike traditional forecasting models such as Box & Jenkins (1976) and Pankratz (1983), which assume that the series studied are generated by linear processes.

When designing an ANN model, we can imagine that it works with a network of artificial "neurons" organized into layers. The variables used to predict (inputs) a dependent variable (output) form the lower layer, while the predicted variables form the upper layer. The ANN model also allows for the possibility of intermediate layers, generally known as hidden layers (Sharda & Patil, 1992)

Designed to represent how the human brain processes information, ANNs are computer algorithms that add knowledge by detecting patterns and correlations and can be trained through experience. They are made up of hundreds of artificial neurons (or nodes) interconnected in hierarchical layers. Each neuron has a specific output function and the connection between each two nodes has a weight, constituting its artificial neural network memory. It is through these weights that the power of neural computations is reflected, i.e. the degree of influence that one cell exerts on another.

Built to simulate the biological function of a neuron, each node has weighted inputs, a transfer function and an output. Feedforward neural networks transmit information linearly, from the input layer to the output, and are the most popular ones used in various applications (Figure 1) (Agatonovic-Kustrin; Beresford, 2000; Gómez, Fernández & Peñuela, 2021).

In this research, the process begins with data entry, in which the explanatory variables correspond to the partial indices generated by factor analysis, and the dependent variable is represented by GHG emissions (Y_i). The data was randomly divided into two sets: 70% was used to train the model and 30% was reserved for the test set (Liu & Cocea, 2017; Dao et al., 2020). The output of a neuron can be written mathematically:

$$Yi = f(n) \quad (12)$$

Where n is the weighted sum of the input signals plus an adjustment term (bias), defined as:

$$n = \sum_{i=1}^{p}(w_j \,.\, X_j) + b \quad (13)$$

Where X_{i1}, X_{i2}, ..., X_{ip} are the neuron's input signals (partial indices generated by factor analysis); w_1, w_2, ..., w_p are the weights associated with each input, determining the importance of each signal in the process; “b” is the bias term, used to adjust the flexibility of the model; and f(*) is the activation function, responsible for the non-linearity of the model, enabling the network to learn complex relationships between the data.

The model's performance was assessed using quantitative metrics, including the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE). The lower the estimated values for these measurements, the better the adjustments. The RMSE is calculated by the root mean square difference between the predicted and observed values, shown in equation 14. It provides an overview of the model's accuracy. The lower the RMSE, the more accurate the model. The Mean Absolute Error (MAE), measured according to equation (15) is also a metric for evaluating models. Finally, the Mean Absolute Percentage Error (MAPE), measured according to equation (16), expresses the errors as a percentage, making it easier to interpret the observed value (Pham et al., 2018; Elsaraiti, 2024). Using several metrics is advantageous for obtaining a broader view of the model's performance from different perspectives (Tripathy & Prusty, 2021).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_{\text{observed}} - v_{predicted})^2} \quad (14)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | v_{observed} - v_{predicted} | \quad (15)$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} | \frac{v_{observed} - v_{predicted}}{v_{predicted}} | * 100 \quad (16)$$

RESULTS AND DISCUSSION

To make the presentation and discussion of the results clearer, they have been organized according to the timeline of the research objectives.

Results found for the first objective

Table 3 show the absolute and relative frequencies of the MATOPIBA municipalities for states that showed an increase or decrease in GHG emissions between 2006 and 2017.

Table 3. Absolute and relative frequencies of MATOPIBA municipalities in relation to GHG emissions between 2006 and 2017

States	Absolute frequencies of municipalities where the ratio of GHG emissions between 2017 and 2006 was less than 1 (GHG < 1)	Relative Frequence (%)	Absolute frequencies of municipalities where the ratio of GHG emissions between 2017 and 2006 was greater than 1 (GHG >1)	Relative Frequence (%)
Maranhão	50	37	85	63
Tocantins	33	23,7	106	76,3
Piauí	8	24,2	25	75,8
Bahia	9	30	21	70
Total	100	100	237	100

Source: Based on the survey results.

It can be seen that there was a predominance of municipalities with an increase in GHG emissions in the period analyzed. Of the 337 municipalities in the MATOPIBA region, 237 (approximately 70.3%) showed an increase in emissions, while 100 (29.7%) registered a reduction. Of the municipalities in the region belonging to the state of Maranhão, 50 out of a total of 135 (37%) showed a reduction in emissions, while 85 (63%) showed an increase. In the state of Tocantins, out of a total of 139, 33 (23.7%) registered a drop in emissions, while 106 (76.3%) showed an increase. In Piauí, out of a total of 33, only 8 municipalities (24.2%) reduced emissions, while 25 (75.8%) showed an increase. In the state of Bahia, out of a total of 30 municipalities in MATOPIBA, 9 (30%) reduced emissions, while 21 (70%) showed an increase (Table 3).

Results obtained for the second objective

Table 4 shows the results found using the principal component decomposition procedure of the factor analysis (FA) used in the research.

Table 4. Results found showing the decomposition of the 11 original variables into 3 main components

		Rotated components Matrix			Component Score Coefficient Matrix		
Variables		1	2	3	1	2	3
Index of rainfall	X_{i1}	0.000	0.009	**1.000**	-0.001	-0.003	**0.500**
Index of vegetation cover	X_{i2}	-0.296	**0.856**	-0.006	0.036	**0.281**	-0.008
Index of agricultural production value	X_{i3}	-0.225	**0.896**	0.013	0.061	**0.306**	0.001
Index of relative livestock value	X_{i4}	-0.035	**0.788**	0.008	0.098	**0.291**	-0.001
Index of rlative recovered areas	X_{i5}	-0.264	**0.941**	0.009	0.056	**0.317**	-0.001
Index or relative cattle quantity	X_{i6}	**0.927**	-0.260	-0.009	**0.219**	0.029	-0.004
Index of_relative_machinefy	X_{i7}	**0.947**	-0.170	-0.024	**0.237**	0.066	-0.013
Index of relative expenditure in pesticides	X_{i8}	**0.865**	-0.258	0.013	**0.202**	0.021	0.007
Index of_relative_industrial GNP	X_{i9}	**0.927**	-0.083	-0.001	**0.243**	0.096	-0.002
Index of relative_cv rainfall	X_{i10}	0.000	0.009	**1.000**	-0.001	-0.003	**0.500**
Index of relative Burn scars areas	X_{i11}	**0.937**	-0.199	0.020	0.230	0.053	0.010
Kaiser-Meyer-Olkin Measure of Sampling Adequacy:							0.738
Bartlett's Test of Sphericity							
Approx. Chi-Square							8186,727
Degrees of Freedom							55
Significance level.							0.000
Total Variance Explained (%)							88.232
Variance Explained by Component 1 (%)							40.464
Variance Explained by Component 2 (%)							29.575
Variance Explained by Component 3 (%)							18.193

*Sources: Results found in the search

Observations: Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization. Component Scores. Rotation converged in 4 iterations.

From the evidence presented in Table 4, it can be seen that the adjustment found using the principal component decomposition (PCD) method was statistically significant. In fact, the Barttlet test, with a high level of significance ($p=0.00$), showed that the correlation matrix of the independent variables used is not an identity. The estimated statistic for the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) test was 0.738, and the total variance explained by the adjusted model was approximately 88.232%. The variances explained by each estimated component after Varimax orthogonal rotation were 40.464%, 19.373% and 18.137%, respectively, for components 1, 2 and 3. These results indicate that the greatest synergy captured by the DCP procedure was in component 1, which measures the ratios of: the number of cattle per hectare; the number of tractors and machinery; spending on pesticides; industrial GDP in relation to the total GDP of the municipalities; and the evolution of burn scars in the period studied. All these variables captured in this component, as assumed in this research, must have positively affected greenhouse gas emissions between 2006 and 2017. This synergy, as we have seen, is responsible for explaining 40.464% of the total explanatory capacity of the model generated (Table 4).

From the results shown in Table 4, it can also be seen that associated with the second component generated in the research, whose variance explains approximately 29.575% of the total explaned variance, are four of the five variables that are supposed to cause a reduction in greenhouse gas emissions: the vegetation cover index; the index that measures the productive potential of crops; the index that measures the productive capacity of animal husbandry and the index that measures the recovery of degraded areas.

For the third component, the greatest synergies were between the variable's rainfall index, which is supposed to reduce GHG emissions, and the index measuring rainfall instability, as measured by the coefficients of variation, which is supposed to increase GHG emissions. It can be seen that these two variables explain 18.193% of the total explained variables. Based on these results, the matrix was generated, which is made up of 3 factor scores that capture these synergies (Table 4).

Results obtained for the third objective

As shown above, the third objective this research sought to assess how GHG emissions are influenced by the synergies between the indicators analyzed. To do this, the GHG emissions ratios between 2006 and 2017 were transformed into indices. In the previous step, when using FA, it was assumed that the relationships between the variables were linear. Thus, the three estimated factor scores are linearly independent. In this step, it is assumed that the relationship between the GHG emissions index between 2006 and 2017 and the independent variables transformed into factor scores is non-linear. The artificial neural network (ANN) model is used to perform the test. The results are shown in Table 5 and Figure 3.

Table 5. Estimation results using ANN, with the GHG emissions index as the dependent variable and the FS1, FS2 and FS3 indices as independent variables

GHG$_{2017/2006}$ emissions ≤ 1		**GHG$_{2017/2006}$ emissions** >1	
Relative error in the Training period (%)	1.011	Relative error in the Training period (%)	1.002
Relative error in the Testing period (%)	0.908	Relative error in the Testing period (%)	0.997
Number of Units in Hidden Layer including 1 bias Hiden Layer	3	Number of Units in Hidden Layer including 1 bias Hiden Layer	4
Independent Variable Importance		Independent Variable Importance	
Variabels	Importance (weights)	Variabels	Importance (weights)
FS1	0.276	FS1	0.486
FS2	0.496	FS2	0.391
FS3	0.228	FS3	0.123
Acuracy tests		Acuracy tests	
RMSE	3.12	RMSE	10.49

MAE	2.52	MAE	6.58
MAPE	13.11	MAPE	21.29

Sources: Results found in the search

The results presented in Table 5 and Figure 3 show that, in the 100 municipalities where there was a reduction in GHG emissions between 2006 and 2017, the score of the factor that brings together the four variables (FS2) that hypothetically impact on this reduction had the highest weight (0.496), followed by the score of FS2, bringing together the variables that, by hypothesis, contributed to the increase in GHG emissions between 2006 and 2017, with a weight of 0.276. Meanwhile, FS3, which brings together the variables associated with rainfall and instability, had a weight of 0.228.

In the case of the 6 variables transformed into indices for which it was hypothesized that they contributed to the increase in GHG emissions between 2006 and 2017, five were grouped into score factor 1 (SF1). The estimated weight was 0.486. Factor score 2 (FS2) had a weight of 0.391 in this definition and factor score F3 had a weight of 0.123 (Table 5).

These results prove the hypotheses of this research, indicating that the variables that were supposed to cause a reduction and increase in GHG emissions between 2006 and 2017 in Matopiba. It can also be seen that the prediction errors in both the training and testing phases were quite low. In addition, the RMSE, MAE and MAPE tests also showed very low values, thus confirming the accuracy of the adjustments.

As shown in the evidence presented in Table 5, to estimate the weights associated with each of the independent (explanatory) variables in the group of 100 MATOPIBA municipalities where the GHG ratio ≤ 1, the ANN model used 3 Units in Hidden Layer, including 1 Hiden Layer bias. To estimate the importance of the explanatory variables in municipalities where the GHG ratio > 1, the model used 4 Units in Hidden Layer, including 1 Hiden Layer bias.

It can be seen that in the case of municipalities where GHG ≤ 1, the highest weight was estimated for the variable FS2 (0.480), the second highest was associated with the variable FS3 (0.286) and the third lowest weight was estimated for the independent variable FS3 (0.234). These results show the relevance of all the original variables that were synthesized in FS2. All of them are supposed to contribute to reducing greenhouse gas emissions. The research hypothesis is therefore confirmed.

With regard to the estimated weights for the variables constructed linearly for the municipalities in which the GHG ratios > 1, it can be seen that, as expected, the greatest weighting was given to the FS1 combination (0.629), which synergistically brings together practically all the original variables that are assumed in this research to have contributed to the increase in greenhouse gas emissions between 2006 and 2017. It can be seen that the rainfall ratios, as well as the rainfall instability ratios measured by the coefficients of variation, both contributed to a reduction and an increase in GHG emissions.

CONCLUSIONS

Based on the research findings, it was observed that of the 327 municipalities within the MATOPIBA agricultural frontier, 100 reduced their Greenhouse Gas (GHG) emissions between 2006 and 2017, whereas the remaining 227 recorded an increase. To analyze the drivers of this phenomenon, the study tested 11 variables—five linked to emission reduction and six to emission increases—using an innovative two-stage methodology. Initially, a factor analysis was conducted to group the variables into factors, followed by the application of artificial neural networks to assess the influence of each factor on GHG variations.

The results confirmed the study's initial hypotheses. The variables hypothesized to cause a decrease in GHG emissions were grouped into a single factor, which carried the most explanatory weight in the municipalities where reductions occurred. Similarly, the variables associated with increased emissions were grouped into another factor that proved most influential in the municipalities where emissions rose. These findings indicate pathways for more sustainable agricultural practices in the region and suggest key variables for future studies on GHG emissions.

REFERENCES

Agatonovic-Kustrin, S.; Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis, v. 22, n. 5, p. 717–727.

Bezerra, F. N. R. (2022). Evaluation of Low Carbon Emission and Climate-Smart Agriculture in Brazil. 265 pages. Doctoral Thesis — Federal University of Ceará, Center for Agricultural Sciences, Graduate Program in Rural Economics, Fortaleza.

Box, G. E. P.; Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Revised ed. San Francisco: Holden-Day.

Brazilian agricultural research corporation (Embrapa). About Embrapa. Available at: https://www.embrapa.br/en/sobre-a-embrapa. Accessed on: Oct. 17, 2024.

Dao, D. V., Adeli, H., LY, H.-B., Le, L. M., Le, V. M., Le, T.-T., & Pham, B. T. (2020). Sensitivity and robustness analysis of GPR and ANN for high-performance concrete compressive strength prediction using a Monte Carlo simulation. Sustainability, 12(3).

Elsaraiti, M. (2024). Short-term power consumption forecasting using neural networks with first- and second-order differencing. Academia Green Energy, v. 1, n. 3, 2024.

Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: General overview. Korean Journal of Radiology, 18(4), 570–584.

Liu, H., Cocea, M. (2017). Semi-random partitioning of data into training and test sets in granular computing context. Granul. Comput. v. 2, p.357–386.

Mapbiomas. RAD 2023: Annual Deforestation Report in Brazil 2023. São Paulo: MapBiomas, 2023. 154 p. Available at: http://alerta.mapbiomas.org. Accessed on: Oct. 28, 2024.

Mapbiomas project. Mapping of burned areas in Brazil between 1985 and 2023 - Collection 3. Available at: https://plataforma.brasil.mapbiomas.org/hotsite-fogo. Accessed on: Oct. 28, 2024.

Noojipady, P., Morton, C. D., Macedo, N. M., Victoria, C. D., Huang, C., Gibbs, K. H., & Bolfe, L. E. (2017). Forest carbon emissions from cropland expansion in the Brazilian Cerrado biome. Environmental Research Letters, 12(2).

Pankratz, A. (1983). Forecasting with Univariate Box-Jenkins Models. New York: Wiley & Sons, Inc.

Pham, B. T., Son, L. H., Hoang, T.-A., Nguyen, D.-M., & Bui, D. T. (2018). Prediction of shear strength of soft soil using machine learning methods. Catena, 166, 181–191.

Rausch, L. L., Gibbs, H. K., Schelly, I., Brandão, A., JR., Morton, D. C., Carneiro, A., Filho, Strassburg, B. B. N., Walker, N. F., Noojipady, P., Barreto, P., Meyer, D., & Soares-filho, B. (2019). Soy expansion in Brazil's Cerrado. Conservation Letters, 12(5), e12671.

Santos, J. F. S.; Naval, L. P. (2022). Soy water footprint and socioeconomic development: An analysis in the new agricultural expansion areas of the Brazilian cerrado (Brazilian savanna). Environmental Development, v. 42, p. 100670.

Sharda, R.; Patil, R. B. (1992). Connectionist approach to time series prediction: an empirical test. Journal of Intelligent Manufacturing, v. 3, p. 317–323.

Tripathy, D. S.; Prusty, B. R. (2021). Forecasting of renewable generation for applications in smart grid power systems. In: TOMAR, Anuradha; KANDARI, Ritu (org.). Advances in Smart Grid Power System. London: Academic Press, p. 265-298

Zhang, G.; Patuwo, B. E, Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, v. 14, p. 35–62.