

## **EVALUATION OF ARTIFICIAL NEURAL NETWORKS AND THE ARIMA MODEL APPLIED TO TEMPERATURE PREDICTION IN A CHARCOAL OVEN**

Rogério Santos Maciel<sup>1</sup>, Nilton Alves Maia<sup>2</sup>, Maurílio José Inácio<sup>2</sup>, Fernando Colen<sup>3</sup>,  
Sidney Pereira<sup>3</sup>, Luiz Henrique de Souza<sup>3</sup>

1 Master student, Program of Computational Modeling and Systems,  
State University of Montes Claros, Brazil.  
rsmaciel@ica.ufmg.br

Corresponding Author

2 Professor, Program of Computational Modeling and Systems,  
State University of Montes Claros, Brazil.

3 Professor, Federal University of Minas Gerais, Brazil.

### **ABSTRACT**

This work used the Artificial Neural Networks RBF- Radial Basis Function and MLP- Multilayer Perceptron, the Neurofuzzy ANFIS- Adaptive-network-based Fuzzy Inference System and the model ARIMA- Auto Regressive Integrated Moving Average, to predict the temperature. Different topologies were tested for Artificial Neural Networks and also for Neurofuzzy, in order to find the best configuration. To evaluate the performance of the predictors, the following metrics were used: Root Mean Square Error - RMSE, Mean Square Error - MSE and Mean Absolute Error - MAE. The analysis of the quality of the predictors was also performed using Theil's U coefficient. Analyzing the results of this study, it was possible to verify that the Artificial Neural Networks, the Neurofuzzy Network and the ARIMA model, can be applied to predict the temperature in charcoal kilns, with the MLP and ANFIS networks presenting the best results.

**KEYWORDS:** Automatic temperature forecast; Methods; Performance.

### **1 INTRODUCTION**

Charcoal is a product resulting from the pyrolysis or carbonization of wood, which is used as fuel in various activities, such as steelworks, steakhouses, domestic uses for stoves and others. One of the areas where coal finds great application is in the production of steel in the steel industry. The good acceptance of charcoal in steel production, where it is used as a reducing agent and fuel for the production of pig iron, comes from its yield, which is superior when compared to mineral coal. In addition, charcoal is a renewable energy source, which brings environmental and competitive advantages to the aforementioned sector, as production from planted areas means that areas of native forest can be preserved, making charcoal production sustainable (FORTALEZA et al., 2019).

In some charcoal kilns, the wood carbonization process is manually controlled. The charcoal burner observes the color and volume of the smoke that comes out of the holes at the base of the

oven, known as armadillos, and in the middle part, known as baianas; the charcoal burner must also observe the presence of embers in the armadillos, and this process must be monitored throughout the furnace heating stage, with interventions being possible day and night.

One way to help the charcoal burner is the implementation of an automatic temperature measurement and forecasting system that helps in the carbonization process and improves working conditions in the charcoal environment. In addition, the temperature forecast through automatic systems prevents people from risking their lives or health in tasks that are carried out in dangerous environments.

The prediction of temperature values also allows for greater speed and precision in the development of activities that are the coal miner's responsibility, thus both models that use computational intelligence techniques, such as Artificial Neural Networks, and those that are based on statistical concepts such as the ARIMA model can be applied for temperature prediction.

In this work, the performance of Artificial Neural Networks, a Neurofuzzy Network and a statistical model for predicting the time series generated from the temperature values collected during the heating process for wood carbonization were evaluated. To perform the prediction, the MLP- Multilayer Perceptron, the RBF- Radial Basis Function, the ANFIS- Adaptive-network-based Fuzzy Inference System and the ARIMA- Auto Regressive Integrated Moving Average model were selected.

## **2 LITERATURE REVIEW**

### **2.1 Time Series**

A time series is a package of data obtained from an analyzed phenomenon, or from the properties of this phenomenon over time by collecting samples at equidistant intervals. Thus, a time series is characterized by the dependence between adjacent observations (CHATFIELD, 2003).

This type of series is of great importance in several areas of science. For example, investors can analyze the future price of shares on stock exchanges, emergency services can predict the possibility of flooding in an area based on historical information collected in previous years (NILSEN, 2021). Health services can monitor the heart rate of patients taking medication to ensure that the heart rate does not change at any time of the day (JOSE, 2022).

Time series can be composed of three types of patterns: trends, cycles and seasonality periods. As a strategy for analyzing a time series, the series can be broken down into these three patterns, making them evident (HYNDMAN; ATHANASOPOULOS, 2018).

Basically, any event that is analyzed sequentially over time is a time series, so time series can be used to study several events, as long as these can be represented by numerical values.

### **2.2 Time Series Prediction**

Predicting is the same as predicting something, that is, it is the way of predicting something that will happen in the future. The prediction can be applied in two situations, when working with very long periods, months or years, or in shorter periods, when there is a need to forecast in minutes or hours ahead.

The prediction can be made by an algorithm that works with rules, or uses statistical probability models. There are three methods used for predicting time series, there is the method based on judgments, where a person's experience is used to predict the future value. Because it does

not use mathematical bases, this method is not widely used. (PARMEZAN, 2016). There is prediction with univariate methods, in which only one type of data is used. Another method is the multivariate method, which uses several different types of data at the same time as the observations; so this time series can be formed by several other time series (REINSEL, 2003).

The prediction of future values of a time series can be obtained in two ways: single-step prediction and multi-step prediction (BROWNLEE, 2018). In single-step forecasting, the forecast is made for a short period of time, immediately after the current one, whereas in multi-step forecasting, medium and long-term horizons are used, in this forecast sets of current values are used for forecasting of a certain instant (BROWNLEE, 2020).

According to Dametto (2018), prediction models with time series can be divided into two categories: computational intelligence techniques and methods based on statistical concepts. It should also be highlighted the existence of methods that combine more than one analysis technique. Among the statistical methods, the following can be mentioned: Exponential Smoothing, Holt Exponential Smoothing, the Regression Method and the Integrated Autoregressive Method of Moving Averages - ARIMA. The methods that use computational intelligence can be divided into two categories: those based on evolutionary computation and those based on Artificial Neural Networks.

### **2.3 Artificial Neural Networks**

According to Guyon (1991), an Artificial Neural Network is a non-linear mathematical structure capable of representing processes, relating inputs and outputs. Artificial Neural Networks can also be described as a structure that uses computers to simulate the functioning of the human brain in a simplified way and that have the ability to identify patterns, treat noisy data, sometimes incomplete and with little precision, in addition to predicting linear systems or not.

Neural networks are considered universal approximation functions, which have more flexible and generalized functions than other statistical methods. Whatever the prediction model, it assumes that there is a hidden relationship between the input and output information. Traditional methods have limitations to determine this relationship, whereas neural networks can be an alternative to work with this type of function (SAÚDE, 2018).

These Artificial Neural Networks can have tens, hundreds and even thousands of processing units, that is, neurons that interact with each other, one transmitting information to another. An Artificial Neural Network is an intelligent structure that resembles the human brain, considering some aspects such as: acquisition of knowledge from the learning process and connection between neurons maintaining the exchange of information (RELVAS, 2009).

In addition to the aforementioned characteristics, Artificial Neural Networks are adaptive models that are easy to train, are also capable of generalizing even in the face of incomplete information, have robust algorithms and can store associative information (JUNIOR, 2010).

#### **2.3.1 MLP-type Artificial Neural Networks**

The MLP Neural Networks - Multilayer Perceptron, make it possible to solve complex problems, which are not possible to be solved by the basic neuron model. These networks, as the name suggests, have input layers, one or more hidden layers which are also known as intermediate layers and output layer. The data presented to the input layer propagates layer by layer passing through the hidden layer to the output layer. Training is done with an algorithm known as

backpropagation, which is based on the rule that corrects the error during training (BARROS, 2018).

MLP networks can be divided into two categories: recursive or recurrent networks and feedforward networks, also known as feedforward networks. In feedforward networks, information is propagated from the input layer, passes through the hidden layers until it reaches the output layer, without any information feedback. Recurrent neural networks have feedback connection in their structure, that is, the outputs of neurons can be connected with neurons from previous layers, or their own input (BAKLACIOGLU, 2021).

These Artificial Neural Networks are widely used to solve problems related to classification and function approximation or regression analysis, which includes forecasting and modeling of time series in several areas (SILVA; SPATTI; FLAUZINO, 2010).

### 2.3.2 RBF- type Artificial Neural Networks

The Artificial Neural Network RBF- Radial Basis Function has only three layers of neurons, an input layer, an intermediate layer or hidden layer that is composed of neurons with radial basis functions and an output layer with neuron with linear activation function, which provides the network response to the pattern applied to the input layer (JÚNIOR, 2021).

An important feature of an RBF Artificial Neural Network is its generalization capacity, which consists of presenting an output suitable for new input patterns, which were not present during the training process to which the network was submitted.

RBF networks have many advantages when compared to Multilayer Perceptron networks, among them we can mention the ability to model any non-linear function with only one hidden layer, another advantage is that with a simple linear transformation in the output layer, the network can optimize using traditional modeling techniques, in addition these networks can be quickly trained.

### 2.3.3 Neurofuzzy Neural Networks

Neurofuzzy Networks are based on a system known as Fuzzy Logic, or Fuzzy Logic. In classical set theory, the elements of a set have a well-defined membership, however in some phenomena it is not possible to clearly define this membership, whether an element belongs or does not belong to a set (MARQUES, 2022).

In 1965 Lotfi Asker Zadeh developed a system that is capable of expressing these phenomena, thus giving rise to Fuzzy Logic. This logic is described by a set of rules of the type: if (antecedent) then (consequent) where the so-called linguistic variables are used in order to define the ranges that the input and output values can assume (SILVA, 2014). Considering as an example, the following situation can be analyzed: if the relative humidity of the air is low and the ambient temperature is high, then the possibility of forest fires is high. This is an example of a rule that can be applied with the help of an expert, but that can be processed using fuzzy logic, as long as numerical values are assigned to the linguistic variables that were used to describe the situation (NOGUEIRA, 2013).

Neurofuzzy networks are capable of combining the semantics of the Fuzzy system based on rules with the ability to learn from Artificial Neural Networks (SOUZA; MAGALHÃES; CAMPOS, 2021).

The Adaptive Neurofuzzy Inference System, ANFIS, is a neural network whose basic idea is the implementation of a Fuzzy inference system through a distributed architecture. This system

created by Roger Jang is successfully applied in function approximation and prediction. The ANFIS network unites the flexibility of Fuzzy systems, offered by the rules, with the architecture of Artificial Neural Networks that provide learning through the use of training data (NOGUEIRA, 2013).

## 2.4 ARIMA model

The Integrated Autoregressive Moving Average (ARIMA) model is based on statistical concepts. This uses three procedures, namely: autoregression, integration and moving averages. The autoregression shows the autocorrelation of the annotations, that is, the influence exerted by an observation on the value of the next one. In the integration it is indicated which number of difference will guarantee the stationarity of the series, and in the moving average are the values that cannot be explained by the past values of the time series. There is also the SARIMA model, which is a generalization of ARIMA, which was developed to work with series with seasonal variation (SANTANA; GONÇALVES, 2022).

The construction of the model takes place through the following steps: Identification, where the model that best describes the behavior of the series is identified; estimation, where the autoregressive component and the parameters are estimated; and the verification step, which assesses whether the estimated model adequately describes the behavior of the data. If the model obtained is not suitable, the identification step is returned and a new cycle is performed.

After performing all the steps, the model that represents the time series data is found, and they can be used to make predictions of future observations.

The combination of the moving average, differentiation and autoregression model methods results in the ARIMA (p, d, q) model, where p is the order of the autoregressive model, d is the degree of differentiation and q is the order of the moving average model. The model can be described by Eq. (1).

$$y_t = \phi_1 y'_{t-1} + \phi_2 y'_{t-2} + \dots + \phi_p y'_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}, \quad \text{Eq. (1)}$$

Where  $y'_t$  is the observation at time t, of the already differentiated time series,  $\phi_i$  is the parameter of the autoregression model, for  $i = 1, 2, 3, \dots, p$ ;  $\theta_j$  is the moving average parameter, with  $j = 1, 2, 3, \dots, q$ ;  $e_t$  represents the process at time t with variance  $\sigma_e^2$  and mean zero.

## 2.5 Cross Validation

Cross-validation is a statistical method of evaluating and comparing learning algorithms. In this method, the ability of the classifier to correctly predict is estimated by dividing the data into training, validation and test subsets (REFAEILZADEH; TANG; LIU, 2009).

The cross validation technique, k fold, is commonly used to verify the generalization power of an Artificial Neural Network, in this technique the training dataset is divided into k subsets (CAMPOS; GARCIA, 2019). It is a fundamental technique when the goal is prediction. This method seeks to estimate the accuracy of a model in practice, that is, what is the performance when a new set of data is applied. In practice, the number of k, which are the partitions, must be 5 or 10, taking into account the computational cost (UFPR, 2022).

When working with prediction models, it is important to guarantee the quality of the model, so that Underfitting or Overfitting does not occur. Thus the sample data must be divided between training, validation and test data.

In the validation stage, only the training and validation data are used, the test data are not used at this time, these must be preserved for the final evaluation of the model. The data is divided into 90% for training and validation and the remaining 10% is used for testing the network. The cross-validation result must be analyzed using model performance metrics such as average, maximum and minimum accuracy, and standard deviation (FIGUEIREDO; OLIVEIRA; SOUTO, 2020).

## 2.6 Performance Metrics

In this topic, the performance metrics used in this work will be presented: the Mean Squared Error, the Root Mean Squared Error and the Mean Absolute Error. The choice of more than one network evaluation method was due to the fact that there was no predominance of one over the other (PARMEZAN, 2016).

The Mean Square Error (MSE) is one of the most used methods to calculate the error and accuracy in neural network models. The calculation is done by summing the variance and the squares of the differences between the predicted values and the actual values. The equation that defines this method is presented in Eq. (2), where  $\hat{y}_i$  is the i-th predicted value,  $y_i$  is the i-th real value and n will be the amount of data used.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{Eq. (2)}$$

The Root Mean Squared Error, or simply RMSE (Root Mean Squared Error), obtained at each simulation, verifies the distance between the predicted data and the correct data. This method was used by Gois et al., (2018) who claim that the lower the value found, the better the network model. The RMSE calculation was performed using Eq. (3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad \text{Eq. (3)}$$

Where  $y_i$  represents the original value,  $\hat{y}_i$  the predicted value and n the number of samples used in the simulation.

The MAE (Mean Absolute Error) is the average of all the absolute errors of the measured variable. Indicates the average size of the error, that is, how many degrees the model misses, regardless of whether the error is negative or positive. Eq. (4) expresses the MAE, where  $y_i$  represents the original absolute value,  $\hat{y}_i$  the predicted value and n the number of samples used in the simulation.

$$MAE = \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{n} \quad \text{Eq. (4)}$$

## 2.7 Coal Furnaces

Coal ovens using baked clay bricks are widely used in Brazil, mainly due to their low cost. These have advantages that make them leaders in the national production of charcoal. Among the advantages are the ease of sealing the walls against the entry of air, which are produced with clayey material similar to that used in the production of bricks, the simplicity in controlling carbonization

and are easy to build, as they do not require an elaborate design of engineering and are easy to maintain (BARCELLOS, 2022).

In this oven model there are armadillos that are openings in the lower part, whose dimensions are 10 x 15 cm. These openings are intended for the admission of atmospheric air during the wood carbonization process. By opening and closing these entries, it is possible to control the internal temperature of the oven, and consequently, the pyrolysis process in it.

### **3 MATERIALS AND METHODS**

To carry out this work, two (2) Artificial Neural Networks were implemented, one MLP and one RBF, one Neurofuzzy ANFIS and one ARIMA predictor. Codes and network configuration were performed using Matlab 2019 software. The work was carried out on a computer that has a 64-bit Windows 10 operating system, AMD A8-6500B APU processor with Radeon (tm) HD Graphics, 3.50 GHz, 4.00 GB of RAM memory installed.

The information used to compose the temperature database was collected in a circular masonry charcoal kiln, installed at the Instituto de Ciências Agrárias of the Federal University of Minas Gerais located at the following geographic coordinates, -16.686215 (latitude), -43.840792 (longitude), in the city of Montes Claros, state of Minas Gerais.

The temperature sensor used in the work was a k-type thermocouple, with a measurement range between 0 and 800 °C. This type of sensor is widely used due to its simplicity and high reliability. The thermocouple is a transducer composed of two wire segments, which are joined at one end. The voltage across the reference junction is called the Seebeck voltage, and it increases as the temperature difference between the measurement junction and the reference junction increases. The thermocouple was connected to a Max6675 temperature reader module with a resolution of 0.25 degrees.

This sensor was installed in accordance with Brasília (2019) which states that the internal temperature of the kiln must be collected in its dome, the point where the values are more constant during carbonization.

After installing the sensor in the appropriate place, it was connected to a Raspberry Pi3 model B board, Quad-core 64-bit arm Cortex A53 processor clocked at 1.2 GHz, 1GB, where the temperature values were recorded. The use of the Raspberry board was due to the fact that it is a complete computer, which has several inputs and outputs, which allows, for example, the connection of a monitor and a keyboard, in addition to the possibility of using a micro SD card for storage. of the operating system and an external SSD or HDD to increase the storage capacity. Figure 1 shows the connection diagram of the circuit, containing the thermocouple, the temperature converter and the Raspberry.

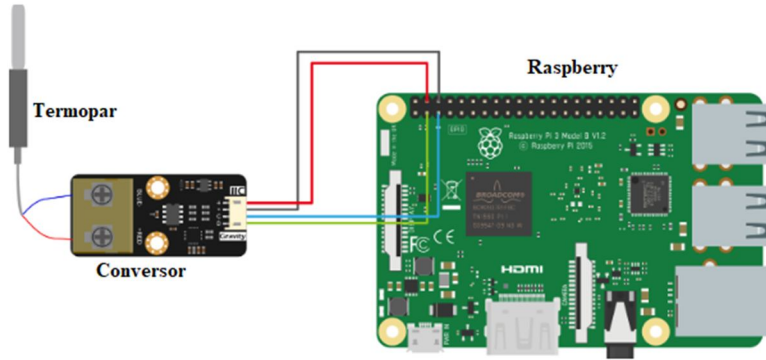


Figure 1. Component connection diagram

While the carbonization process was taking place, 956 samples were collected containing the temperature values recorded every 5 minutes, thus covering the entire heating stage of the wood carbonization process.

### 3.1 Configuration and Testing of Predictors

To define the topology of Artificial Neural Networks and Neurofuzzy, Cross Validation was used. This step was performed as follows:

The data (956 samples) were split into: 90% (860 samples) were set aside to train and validate the networks, and the remaining 10% (96 samples) were used to test the networks. Then the 90% were divided into 5 subsets containing 172 samples each. These 172 samples were separated into 70% (120 samples) for training and 30% (52 samples) for network validation.

After dividing the samples, 3 configurations were defined for each Artificial Neural Network MLP and RBF, and also for the Neurofuzzy ANFIS, as described below:

- MLP network: tested with logsig activation function, a hidden layer and an output layer with linear activation function, in this network the number of neurons in the hidden layer was changed, the other parameters were kept the same in the three configurations.
- RBF network: tested with goal = 1e-3, spread = 1.5, in this network the number of neurons in the hidden layer was changed.
- Neurofuzzy ANFIS: tested with GridPartition grouping method and 2 membership functions, which generates 16 rules. Tested with the SubtractiveClustering grouping method when the chosen parameter was ClusterInfluenceRange = 0.5, and with the FCMClustering method, the NumClusters parameter = 4.

According to the average values of the calculated metrics, it was possible to verify that the Networks with the best configurations were the Neurofuzzy ANFIS with the 'GridPartition' grouping method, the MLP with 1 hidden layer with 5 neurons and the RBF with 5 neurons in the layer hidden. These Networks were configured with window equal to 4 and step equal to 1, in the ARIMA model (1,1,1), the number of lags used was equal to 4, in order to maintain the same criterion for all predictors. The ARIMA predictor, like the others, was configured using matlab software. Within the software, the Econometric Modeler tool was used, with this tool it is possible to import the database and simulate the proposed model.



### 4 RESULTS AND DISCUSSION

When working with time series it is important to determine whether or not the series is stationary. For this, the Excel spreadsheet editor was used. Next, a graph of the time series and the moving average was generated, which are shown in Figure 2, where it is possible to observe that the moving average suffers variations over time, which characterizes the series as non-stationary. In addition to calculating the moving average, the variance of the observed values was also calculated, which were plotted in the graphs referring to these values, as shown in Figure 3.

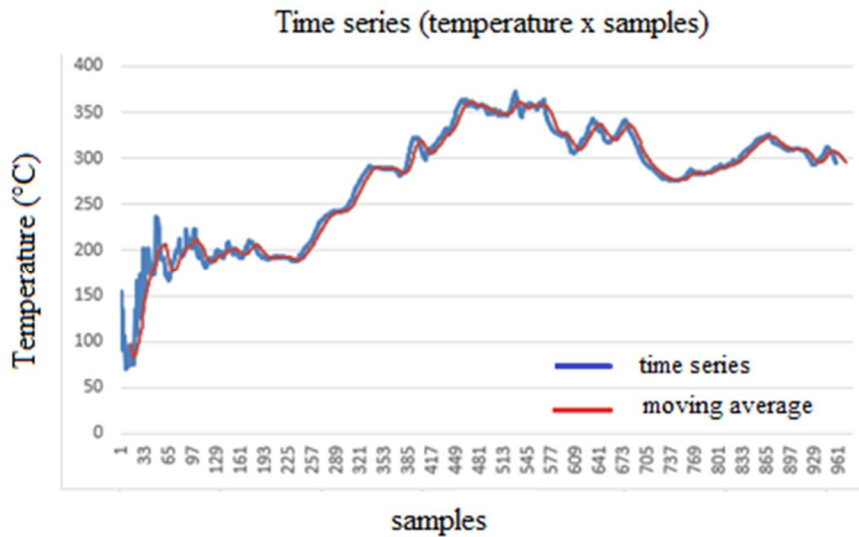


Figure 2. Generated time series

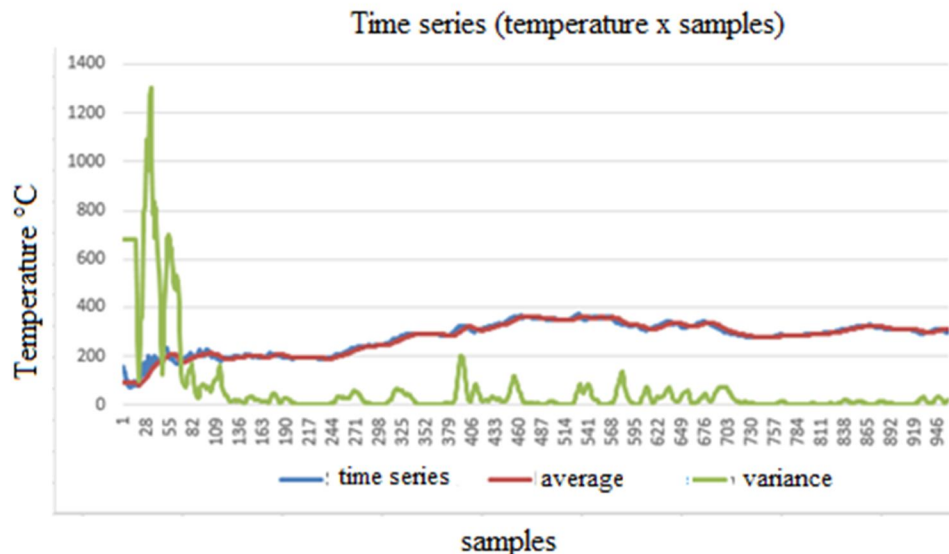


Figure 3. Average and variance of the series

Analyzing the graph, it is possible to verify that the variance, as well as the moving average, is not constant, once again characterizing the series as non-stationary. Figure 4 shows the

decomposed Time Series. In the decomposition, it is possible to notice the time series entitled Model Fit and the presence of the residue, entitled Residual Plot, which is analyzed later.

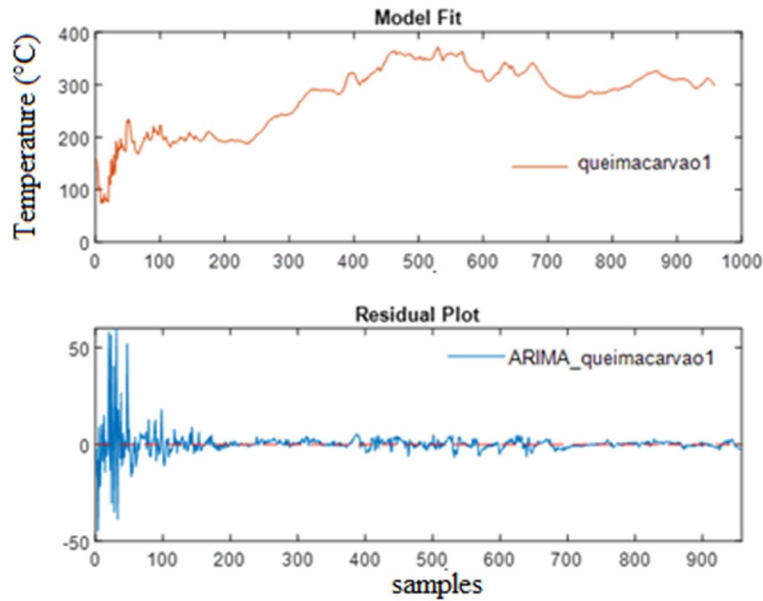


Figure 4. Decomposed Time Series

With the aid of the Matlab software, the ACF graph of the time series was generated, which is shown in Figure 5. In this graph, it is possible to notice that the autocorrelation persists assuming high values, which is characteristic of non-stationary time series.

For the correct determination of the stationarity or non-stationarity of the series, in addition to the graphical analysis, the Augmented Dickey Fuller test was applied. The unit root test was performed using the Matlab adftest function. When the series was loaded, the returned value was zero (0), which confirms the graphical analysis performed and indicates the non-stationarity of the series.

An important step during time series analysis is to verify that the model adequately fits the data. Figure 6 shows the correlogram of residuals for the series under study. It is possible to observe that the errors are not correlated, which indicates a good fit of the model to the data.

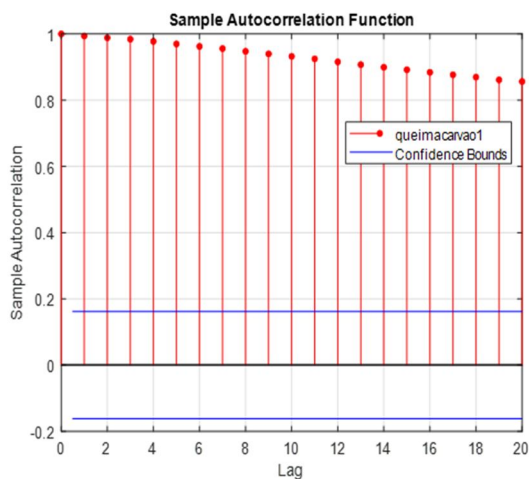


Figure 5. Series Correlation

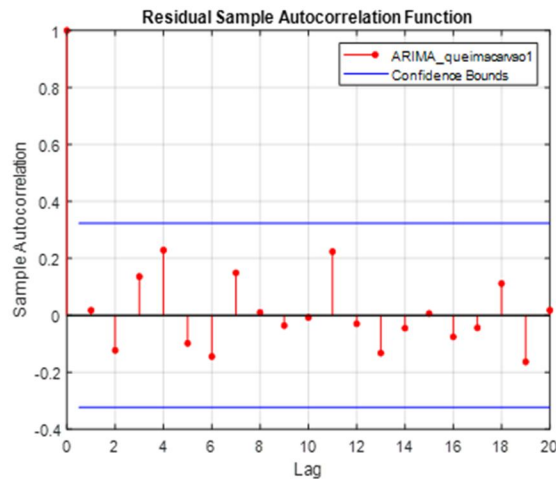


Figure 6. Residual ACF

The residuals can be seen in the histogram of Figure 7. In this, the residuals of the model are concentrated close to zero, showing the good adequacy of the model to the data. The QQ-plot graph in Figure 8 shows that most of the residuals are on top of the line, which demonstrates a normal behavior, with only a few deviations in the tail. In addition to the graphical analysis of the residuals, it was verified whether the  $p$ -value  $> 0.05$ , in this case the  $p$ -value of the ARIMA model (1,1,1) was 0.135, indicating that the model passed the normality test of the residuals.

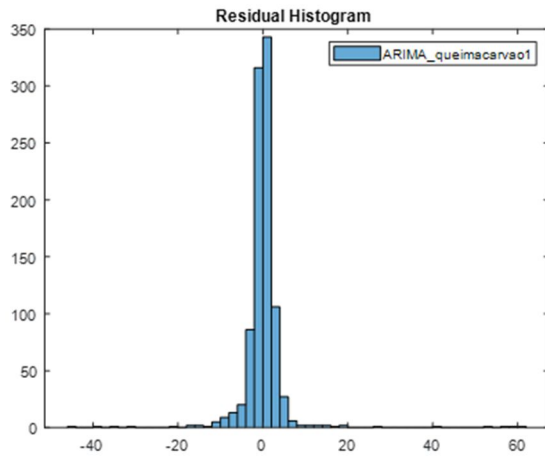


Figure7. Residual Histogram

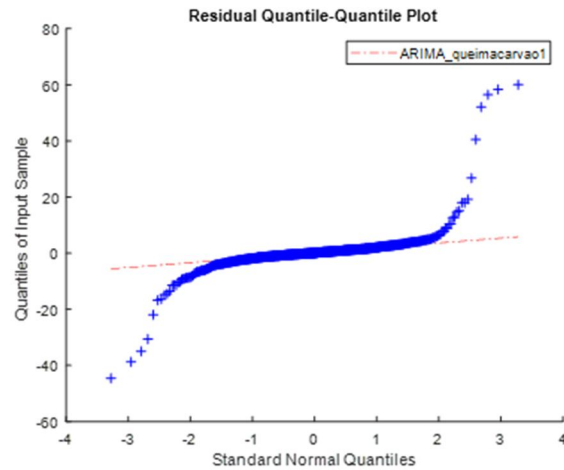


Figure 8. Residual QQ-plot

During the simulation stage, graphs were generated referring to the actual values and predicted values in the three Network models and also for the ARIMA model. These charts are shown in the Figures below.

Figure 9 shows the graph referring to the prediction generated by the MLP Artificial Neural Network. In this figure it is possible to observe the approximation between the real values and the values predicted by the MLP Artificial Neural Network, the proximity is confirmed by the superimposition of the generated graphs, not being possible to distinguish one from the other.

Figure 10 shows the graphs generated by the RBF Artificial Neural Network during training with 100% of the samples. In this Figure, it is possible to notice that in some moments the RBF Artificial Neural Network cannot correctly predict the temperature value based on the real value, red graph. This is evident, for example, when looking at sample 300, where the temperature is approximately 275 °C, at that moment the network indicates a different value between the real and the predicted one.

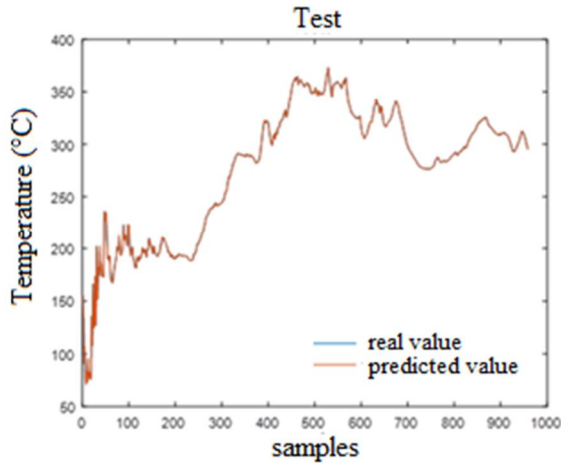


Figure 9. MLP Network Prediction

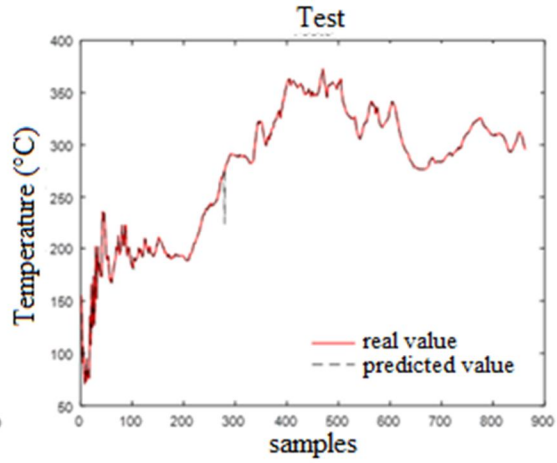


Figure 10. RBF Network Prediction

Figure 11 shows the graph referring to the actual values and the values predicted by Neurofuzzy ANFIS. By analyzing the graph referring to Neurofuzzy ANFIS, it can be seen that there is an approximation between the actual temperature values and the values predicted by the network. This is possible due to the overlapping of the graph referring to the Predicted Value, in red, to the Actual Value graph, in blue.

After analyzing the predictors based on Computational Intelligence, the analysis of the prediction performed by the ARIMA model was performed. In the graphs in Figure 12, which refer to the ARIMA model, it is possible to notice the divergence between the actual value and the predicted value. This difference between the graphs matches the error values calculated by the performance metrics for the model.

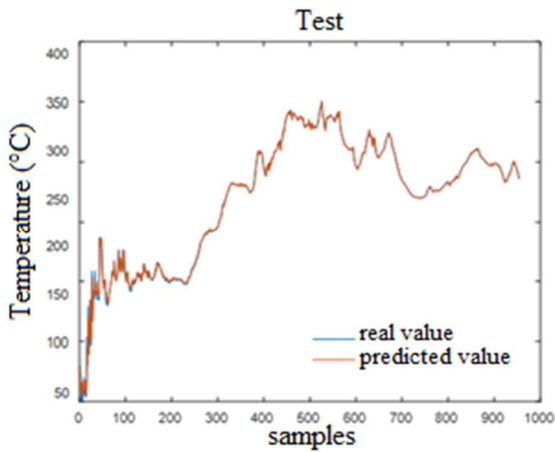


Figure 11. Prediction of the ANFIS Network

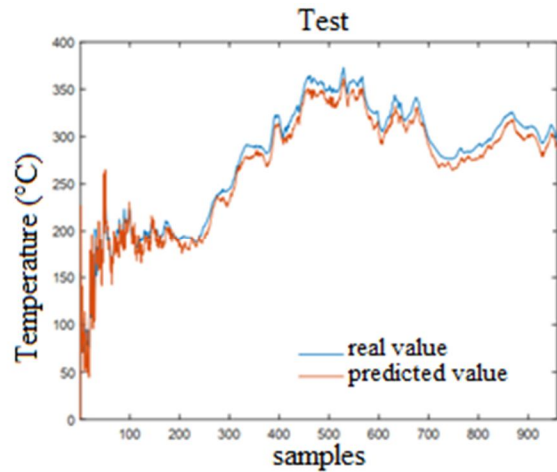


Figure 12. Prediction of the ARIMA

During the simulation stage, performance metrics were calculated to verify which of the predictors best suited the collected data. The following tables show the results of metrics for all predictors when they were tested with 100% of the samples. In Table 1, the MSE of all predictors were registered.

Table 1. MSE of predictors

Ciclo	network MLP 1	network RBF 1	network ANFIS 1	ARIMA
	MSE	MSE	MSE	MSE
1	4,0371e-09	5,9316	0,0123	8,4800

When analyzing Table 1, it is possible to notice that the ARIMA predictor was the one that presented the highest error value. It is also possible to verify that the MLP Artificial Neural Network obtained the best performance. In Table 2, the RMSE of the predictors was recorded.

Table 2. RMSE of predictors

Ciclo	network MLP 1	network RBF 1	network ANFIS 1	ARIMA
	RMSE	RMSE	RMSE	RMSE
1	6,3538e-05	2,4355	0,0113	2,9120

When the calculated metric was the RMSE, the MLP Neural Network obtained the best performance. The ARIMA predictor had the worst performance. The other metric chosen for this work was the MAE, and the results regarding the predictors are recorded in Table 3.

Table 3. MAE of predictors

Ciclo	network MLP 1	network RBF 1	network ANFIS 1	ARIMA
	MAE	MAE	MAE	MAE
1	3,4819e-05	6,8812	0,0052	4,5500

When analyzing the data in Table 3, it is possible to verify that the ARIMA model improved its performance when compared to the other predictors, however the MLP predictor continues to be the best suited in the temperature prediction task.

## 5 CONCLUSION

After conducting this study, it was concluded that Artificial Neural Networks, Neurofuzzy and the ARIMA predictor can be valuable tools when applied to predict temperatures in charcoal kilns. This hypothesis is supported by the low values of MSE, RMSE and MAE calculated during the tests of the predictors, and also by the analysis of the graphs generated by the Matlab software, which show the proximity between the data that are informed to the network and the data generated by the output. Regarding the ARIMA predictor, it presented the worst result for this application, according to the metrics chosen for its evaluation and comparison with the other methods applied in this work.

Thus, it is possible to conclude that the MLP-type Artificial Neural Network was the best fit for this application, being able to contribute significantly to the temperature control process in charcoal kilns. However, it can also be stated that the Neurofuzzy ANFIS cannot be discarded for this application, as it also presented good results during the simulation.

**REFERENCES**

- [1] BAKLACIOGLU, T. (2021). Predicting the fuel flow rate of commercial aircraft via multilayer perceptron, radial basis function and ANFIS artificial neural network. *The Aeronautical Journal, Turkey*, vol. 125, n. 1285.
- [2] BARCELLOS, D. Types of Charcoal Furnaces. Available in: <https://danielbarcellos.com/tipos-de-fornos-de-carvao-vegetal/>. Accessed on December 14, 2022.
- [3] BARROS, V.P.A. (2018). Performance Evaluation of Backpropagation Algorithms with Artificial Neural Networks for Solving Nonlinear Problems. Master's Dissertation, UTFPR, Ponta Grossa.
- [4] BRASÍLIA. (2019). Sustainable Steelmaking. Sustainable Coal Production. Oven-Furnace System, Operation Manual. Federal District.
- [5] BROWNLEE, J. (2018). Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery.
- [6] BROWNLEE, J. (2020). 4 Strategies for Multi-Step Time Series Forecasting. Machine Learning Mastery.
- [7] CAMPOS, L.M.L.; GARCIA, J.C.P. (2019). Management and Technology Magazine, Pedro Leopoldo, vol. 19 no. 1,
- [8] CHATFIELD, C. (2013). The analysis of series: An introduction. Boca Raton, United States of America: Taylor e Francis, (Chapman and Hall/CRC).
- [9] DAMETTO, R.C. (2018). Study of the application of artificial neural networks for the prediction of financial time series, Master's Dissertation, UNESP, Bauru, 2018.
- [10] FORTALEZA, A.P. et al. (2019). Biomass of Forest Species for the Production of Charcoal. *Ci. Fl.*, Santa Maria, vol. 29, no. 3.
- [11] FIGUEIREDO, A.F.; OLIVEIRA, H.S.; SOUTO, E.J.P. (2020). Dense Neural Networks for Stress Classification. *J. Health Inform. SBIS Special Number*.
- [12] GOIS, G.A. et al. (2018). Artificial Neural Networks for Predicting the Total Fuel Consumption of a Blast Furnace. Brazilian Association of Metallurgy, São Paulo
- [13] GUYON, I. (1991). Neural networks and applications tutorial. *Physics Reports*. vol. 17, n. 3.
- [14] HYNDMAN, R.J.; ATHANASOPOULOS, G. (2018). Forecasting: principles and practice, 2nd edition, O Texts: Melbourne.

- [15] JOSE, J. (2022). Introduction to Time Series Analysis and its Applications. Christ University, Bangalore.
- [16] JÚNIOR, A.C.S. (2021). Prediction and Importance of Predictors in Approaches Based on Computational Intelligence and Machine Learning. Thesis, Doctorate in Genetics and Breeding, UFV, Viçosa.
- [17] JUNIOR, A.D.A. (2010). Nonlinear Prediction of Oil Production Curves via Recursive Neural Networks, Master's Dissertation, UFRN, Natal.
- [18] MARQUES, C.M. (2022). Neuro-Fuzzy Model of Customer-Based Decision Support in a Shopping Center. Rev. Lat.-Am. Innovation Eng. Product Curitiba, Paraná, Brazil, no. 10.
- [19] NILSEN, A. (2021). Practical analysis of time series: prediction with statistics and machine learning. Rio de Janeiro: Alta Books.
- [20] NOGUEIRA, M.M. (2013). Applying Fuzzy Logic in the Control of Mobile Robots using Programmable Logic Devices and the VHDL Language, Master's Dissertation in Electrical Engineering, UNESP, Ilha Solteira.
- [21] PARMEZAN, A.R.S. (2016). Prediction of time series by similarity, Master's Dissertation in Computer Science and Computational Mathematics, USP, São Carlos.
- [22] REFAEILZADEH, P.; TANG, L.; LIU, H. (2019). Cross-Validation. In: Encyclopedia of Database Systems. Boston, MA: Springer US.
- [23] RELVAS, M.P. (2009). Biological Foundations of Education: Awakening Intelligences and Affectivity in the Learning Process. 2nd ed. Rio de Janeiro: Wak.
- [24] REINSEL, G. C. (2003). Elements of multivariate time series analysis. New York, United States of America: Springer Series in Statistics, 2 ed.
- [25] SAÚDE, L.M.S. (2018). Comparative analysis between autoregressive methods, integrated moving averages and artificial neural network for forecasting time series, Master's Dissertation in Production Engineering, UFI, Itajubá.
- [26] SANTANA, M.V.L.; GONÇALVES, E.W.S. (2022). Application of the ARIMA model for forecasting sales of the Rossmann pharmacy chain.
- [27] SILVA, I. N.; SPATI, D. H.; FLAUZINO, R. A. (2010). Artificial Neural Networks: for engineering and applied sciences. Sao Paulo: Artliber.
- [28] SILVA, I.A.F. (2014). Applications of Neural and Neurofuzzy Networks in Biomedical Engineering and Agronomy, Doctoral Thesis in Electrical Engineering, UNESP, Ilha Solteira.

[29] SOUZA, R. M. de; MAGALHÃES, R. R.; CAMPOS, A. T. (2021). Neuro-fuzzy model for predicting the energy input of different concrete dosages in buildings. *Built Environment*, Porto Alegre, vol. 21, no. Two.

[30] UFPR. Sampling Methods. Available in:

<http://cursos.leg.ufpr.br/ML4all/apoio/reamostragem.html>. Accessed on December 18, 2022.