APPLICATION OF THE NONLINEAR AUTOREGRESSIVE MODEL WITH EXOGENOUS INPUTS FOR RIVER LEVEL FORECAST IN THE AMAZON

Gisele de Freitas Lopes

gikalps@gmail.com

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

Manoel Henrique Reis Nascimento

hreys@itegam.org.br

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

Alexandra Amaro de Lima

xanduca@gmail.com

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

Nadime Mustafa Moraes

nmoraes@uea.edu.br

Academic Department of the Amazon State University - UEA

José Roberto Lira Pinto Júnior

robertojunior72@gmail.com

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

Ana Priscila Barbosa de Alencar

wertprisca@gmail.com

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

David Barbosa de Alencar

david002870@hotmail.com

Academic, of the Postgraduate Program in Engineering, Process Management, Systems and Environmental (PPEPMSE) - Institute of Technology and Education Galileo of the Amazon – ITEGAM.

Abstract

The present work is justified by three basic lines that involve the problem of the theme, which are the use of Artificial Intelligence, the problem of floods in the Amazon and the issue of technology in favor of decision making. The environmental impacts caused by economic and social factors are problems portrayed in scenarios such as floods and ebbs of rivers, bringing up situations such as an increase in diseases, reduction of agricultural production in locations that depend on accurate geological control, in addition to the increase in erosive processes. in risk locations. Thus, the use of AI to predict the river level, which consequently can minimize problems arising from floods that cause an environmental impact, is highly possible, since when it is known in advance that an event is close to happening, decisions can be taken so that the impacts be smaller. This work models and applies NARX to forecast the river level in the Amazon with variables of easy access and implementation through the MATLAB software, in order to contribute with a forecast model capable of predicting a possible flood from the river level.

Keywords: Forecast; river level; NARX, Artificial Intelligence.

1. Introduction

In Brazil, natural floods occur every year and some caused over the years, causing damage and these could be avoided if they were foreseen. There are a variety of methods for forecasting river-level floods that can help society, especially riverside people. For stochastic flood prevention, a large amount of data samples is needed, so that the accuracy of the forecast is high and thus can be effective (FAVA, 2015).

In the North of Brazil, there is a large number of heavy rains and these events cause flooding and possibly economic and social losses, such as loss of materials and individuals, proliferation of diseases and homeless people (SILVA, 2006).

Among the existing forecast models, the stochastic ones stand out, which use data, preferably in real time, to make the forecast with more probability of success. Among the stochastic models, there are those that use Artificial Intelligence techniques, which are used because they have greater feasibility of implementation and satisfactory accuracy, which provides a greater potential for flood alerts, in order to avoid catastrophes and large losses (FAVA). , 2015).

Stochastic models that make use of Artificial Neural Networks have the advantage of not needing explicit knowledge of the watershed, since ANNs work well in this scenario due to the way in which the neurons that compose them are learned (MARACAJÁ, 2005; GORODETSKAYA, DA FONSECA and DE MELO RIBEIRO, 2017).

The objective of this work is to apply the non-linear autoregressive model with exogenous inputs to forecast the river level in light of the scenarios of environmental impacts caused by floods or ebbs in the state of Amazonas.

To achieve this objective, the following steps were carried out: Identify the significant input variables with correlation with the output variable for river level prediction, define optimal settings (number of layers, training algorithm, activation functions, number of neurons). for application in the forecast model, simulate forecasts with the ideal configurations of the model to verify its accuracy, compare the results obtained

through NARX with the SARIMA model, in order to verify the forecast performance with another time series forecasting algorithms.

2. LITERATURE REVIEW

2.1 Amazon water resources

A "Water Resource" can be understood as any and all surface and groundwater available for use, in turn, employed in a particular use or activity (DE AMORIM, 2021).

This resource gives life to several sources of technologies and economic goods, as a result of which the National Water Agency (ANA) was created with the objective of managing these water resources and regulating access to water, in such a way as to allow sustainable use. these resources for the benefit of the current and future generation.

However, Law No. 3,167, of August 28, 2007, reformulates the disciplinary rules of the State Water Resources Policy and the State Water Resources Management System (SINGREH).

In 1997, Brazilian water resources began to be regulated by Federal Law nº 9.433/97, this law represents a fundamental milestone in the process of changing the institutional environment that regulates water resources in Brazil by instituting the National Policy on Water Resources and creating the System National Water Resources Management (SNGRH) (SILVA, MIRANDA and SANTANA, 2016).

In the Amazon, the river transport network is one of the most used means due to the dimension and proportion of water resources found in the Amazon basin, located in Brazil and in seven other South American countries, namely: Bolivia, Colombia, Guyana, French Guiana, Peru, Suriname and Venezuela (NATTRODT and DIAS, 2021), enjoys approximately 7 million square kilometers in length, 4 million of which in Brazilian territory, in addition to 23 thousand kilometers of navigable rivers, constituting rivers such as: Negro, Solimões , Branco, Juruá, Xingu, Japurá and others. Silva et. al., (2013) illustrates the size of the Amazon River Basin.

## 2.2 Overview of the flood and elower of the Amazonas river

The Amazon River is located in the largest hydrographic basin on the planet, starting in the Andes Mountains, covering a distance of 6,577 km until reaching its mouth in the Atlantic Ocean, with the contribution of more than a thousand tributaries on both the left and right banks (ALBUQUERQUE, 2018).

In 2021, the Amazonian rivers presented quotas between the daily maximums since the beginning of the year and from the month of April, there were severe floods along the basin, following the prognosis of the Geological Service of Brazil (SGB-CPRM). The peak of the flood in the Amazon River basin is normally observed between the months of June and July (CPRM, 2021).

What determines the magnitude of these floods are the rains that occur in all the basins that drain into this region, such as the Negro and Solimões basins and all their tributaries (Purus, Juruá, Japurá, Jutaí, etc.), including their areas external to Brazil, in Colombia, Peru and Ecuador. The Amazon River Basin is the largest watershed in the world (ALVES, 2021; CPRM, 2021).

The Graph in Figure 1 presents Cotagram data for the Manaus (Rio Negro) station for the year 2021, in which it is possible to identify the daily maximums recorded, daily minimums recorded, the minimum

observed record of the 2010 historical series, the median and the permanence series from 2021 to 2021, the envelope curves represented by the blue band characterize data between 15 and 85% permanence for daily quota data (CPRM, 2021).

Figure 1 - Cotagram of the Rio Negro in Manaus.
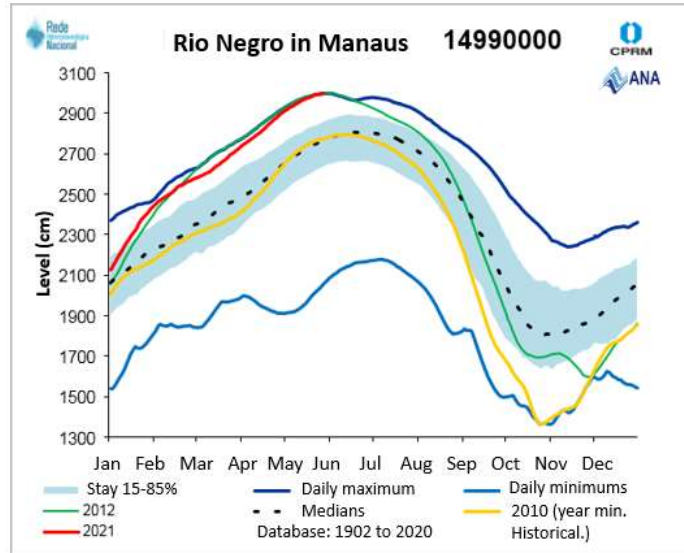
Source: (ALVES, 2021; CPRM, 2021).

According to Alves (2021) the blue band curves presented in Figure 1 express values to be analyzed carefully, that is, for values above the band, it represents an expressive flood process and, for the values below, an accentuated ebb process.

Figure 2 illustrates a comparison of historical highs and lows recorded in Manaus, capital of the state of Amazonas, over a period from 1903 to 2019.

Figure 2 - Maximums and minimums in the period from 1903 to 2019.
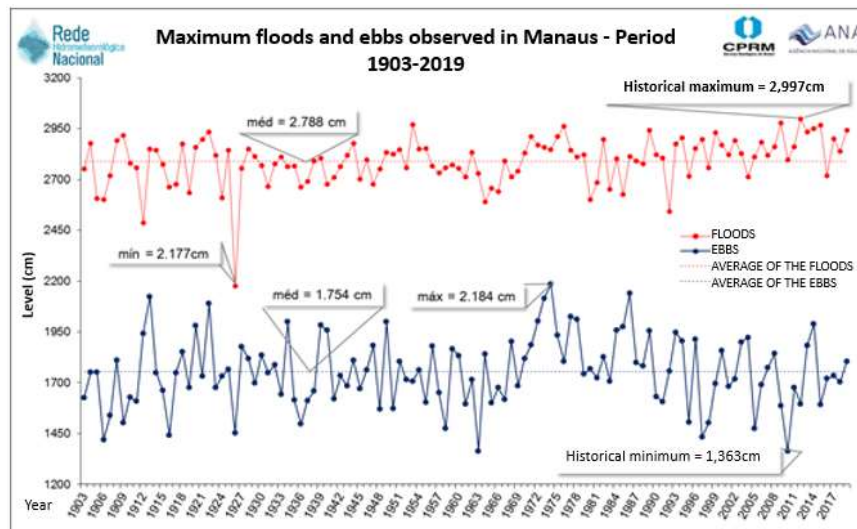
Source: (ALVES, 2021; CPRM, 2021).

It is possible to identify that the historical maximum of 2.997 cm is reached in 2012 as shown in Figure 2, this shows the expressive increase in the river level as a result of external events that directly influence the change in hydrological behavior. Every year, global climate influences affect and harm the situation of the residents of the riverbanks and the presence of deforestation further aggravates the situation of climate change within the Amazon.

2.3 Methods and models for time series prediction

In the literature there are some models for forecasting time series, among them the following models can be cited: ARIMA, SARIMA, RNA and NARX.

ARIMA is generally used for short-term forecasts, in which the data behavior is not seasonal, whereas SARIMA is used for series that have seasonality. ANNs and the NARX model can be used for both types of time series (BLÁZQUEZ-GARCÍA et al., 2020; WANG et al., 2015).

2.3.1 Statistical Models

Statistical models are simple to implement and have the characteristic of having a good performance when the forecast is short-term, these models use a historical data to base and predict steps forward (CHANG et al., 2014).

Statistical models use time series methods such as: Auto Regressive - AR, Auto Regressive Moving Average - ARMA, Auto Regressive Integrated Moving Average - ARIMA and Seasonal Auto Regressive Integrated Moving Average (SARIMA) (ALENCAR et al., 2018).

Statistical models also address computational methods such as: RNA, NARX and Neuro-Fuzzy (BARBOSA DE ALENCAR et al., 2017).

2.3.2 River Level Forecast Models

Among the models used for river level prediction, there are the following: RNA, Hydrological Models such as Moving Difference Filter (FDM), Exponential Moving Average Filter (FMME) and Gamma Moving

Average Filter (FMMG).

ANNs are commonly used for time series forecasting, but for the prediction of river and flood levels there is a very limited amount, in view of this, the study carried out by (ALBERTON et al., 2021) used LSTM type ANN (Long Short-Term Memory) and MLP (Multi-Layer Perceptron) for forecasting the Itajaí-Açu river in Blumenau.

It is possible to observe some works that used RNA to predict the river level.

(ALBERTON et al., 2021) Application of artificial neural networks to forecast floods in the Itajaí-Açu river in Blumenau, SC, Brazil, LSTM and MLP RNA method, The study makes a comparison between an LSTM and MLP type ANN and proposes the use of LSTM-type ANN to forecast river level and floods.

(CRISTALDO et al., 2020) Artificial Neural Networks applied to flood forecasting for the Pantanal region in Mato Grosso do Sul, RNA method of the MLP type. This research used the RNA technique to predict floods in the Aquidauana River.

(ARAÚJO et al., 2020) Seasonal Flow Forecast for the Orós Basin (Ceará, Brazil) Using Neural Networks and the K-Neighbors Resampling Technique, RNA Method and K-Neighbors Resampling, The article compares two techniques (RNA and Resampling of K-neighbors), applied to the Orós dam, which is located in the Alto Jaguaribe Basin, where the RNA technique had better results.

(FINCK, 2020) Current-time forecast of fluvial levels with Artificial Neural Networks: Application to the Taquari-Antas/RS river basin, Hydrological models with RNA N, this study was made a hybrid method combining a hydrological model (FMMG) with RNA for forecasting floods in the Taquari-Antas basin.

2.4 RNA Model

The ANNs were initially constituted in 1943 by Mcculloch and Pitts, in which this AI technique was based on the functioning of human neurons, thus imitating the learning capacity that the brain has to learn, having as a starting point a data history that simulates human experience over time and thus can determine a prediction (MCCULLOCH and PITTS, 1943).

According to Haykin (2008), the neural network is like a processor with numerous processing units acting at the same time, which has the ability to store data and show the knowledge acquired through data processing.

The representation of the human neuron and the Artificial Neural Network (ANN) can be exemplified in CERRI (2020) and PARENTE (2021).

The artificial neuron imitates the behavior of biological neurons in order to learn according to training based on a history of data, and the model of a neural network consists of inputs that can assume any positive value indicating the number of variables to be used in the process. model. ANN models have three layers, being divided into input layer, hidden layer and output layer (SILVA et. al., 2020; DE MENDONÇA et. al., 2021).

Input Layer: In the input layer are the input variables, which are used to achieve the ultimate goal, be it predicting a variable x or y. Since the input variables must have a correlation with the output variable, the greater the correlation between the variables, the better the final result. It is worth mentioning that the forecast loses reliability in the measure of predicted steps ahead, example: Forecasting data from one day ahead is easier than forecasting from 365 days ahead.

Hidden Layer: The hidden layer contains the neurons that are responsible for approximating the most suitable values to achieve the intended final result, such approximation values are known as weights. The hidden layer can be composed of several sublayers and with n neurons in each of them.

Output Layer: The output layer is the final result of the processing done by the Neural Network which may contain one or more dependent variables as a result.

In DE MENDONÇA et al., (2021), the architecture of a multilayer RNA, also known as MLP, is presented. The process of building an ANN has some parameters, among them, there is the definition of the number of hidden layers, the number of neurons in each layer, the definition of the variables to be used in the problem to be solved, defining the algorithm of Neural Network training, find the best transfer function and use parameters to measure ANN performance. These are the main elements for the development of an ANN.

Definition of variables: The first process is to define the input variables, which must have a correlation with the output variable, which is the result to be found. For a good variable definition, there must be the same number of records between both input variables.

Number of hidden layers: The number of hidden layers is usually defined by trial and error tests to arrive at the number that best suits the problem under study. Few hidden layers can harm the performance of the ANN, many hidden layers also happens the same, so it is necessary to test several layers to verify how many layers the studied problem has a better performance.

Number of neurons: With neurons, the premise is the same as the number of hidden layers, it must be tested to find the best amount, thus aiming to achieve the best possible performance of the ANN.

Training Algorithm: The training algorithm is intended to make the ANN learn from historical data. Each algorithm has its particularities, some are better for forecasting seasonal data, others linear, and others for pattern identification. Therefore, it is interesting that in the process of choosing the training algorithm, some algorithms are tested to identify the best one for the problem under study. There are some training algorithms, among the best known can be listed 12 algorithms, namely:

- Levenberg-Marquardt,
- Bayesian Regularization,
- Broyden–Fletcher–Goldfarb–Shanno Quasi-Newton,
- Resilient Backpropagation,
- Scaled Conjugate Gradient,
- Conjugate Gradient with Powell/Beale Restarts,
- Fletcher-Powell Conjugate Gradient,
- Polak-Ribiére Conjugate Gradient,
- One Step Secant,
- Variable Learning Rate Gradient Descent,
- Gradient Descent with Momentum;
- Gradient Descent.

Transfer function: The transfer function is present both in the hidden layers and in the output layer, there are some transfer functions that can be found in REIS et al., (2018).

Performance parameters: There are some performance parameters that can be used to evaluate the performance of an ANN, the most used being MSE, RMSE, Regression and MAPE.

There are some RNA variants, including NARX, which is the technique selected for this study. This model is widely used for time series forecasting (MATKOVSKYY and BOURAOUI, 2019; WUNSCH, LIESCH and BRODA, 2018).

DE MENDONÇA et. al., (2021), shows the operation of the NARX Neural Network, which has an input layer, a hidden layer and an output layer, the main difference between an Artificial Neural Network and NARX is that the NARX model has feedback, being indicated for time series forecasting, especially if there is seasonality in the input and target data.

3. MATERIALS AND METHODS

3.1 Materials

The materials of this article will be presented below.

3.1.1 Significant Variables for the Forecast Model

The variables used for the forecast model are shown in Table 1.

Table 1 - Significant variables for the model.

Variable	Type
Rain	Input
River temperature	Input
Relative humidity	Input
River level	Exit

Source: (AUTHOR, 2021).

3.1.2 Training Algorithms

The training algorithms for the NARX model were of paramount importance to find the best convergence state based on the number of neurons, layers, activation functions and errors such as MSE, RMSE, NRMSE and MAPE. Table 2 presents the algorithms that will be used, characterizing their convergence methods.

Table 2 - Training algorithms.

Algorithm	Method
Levenberg-Marquardt	optimization
Bayesian Regularization	optimization
Broyden–Fletcher–Goldfarb–Shanno Quasi- Newton	Iterative
Resilient Backpropagation	Heuristic
Scaled Conjugate Gradient	Iterative
Conjugate Gradient with Powell/Beale Restarts	Iterative
Fletcher-Powell Conjugate Gradient	Quasi-Newton

Polak-Ribiére Conjugate Gradient	optimization
One Step Secant	Newton
Variable Learning Rate Gradient Descent	Numeric
Gradient Descent with Momentum	Iterative
Gradient Descent	Iterative

Source: (AUTHOR, 2021).

3.1.3 Activation Functions

The transfer functions are of paramount importance to evaluate and reduce the explosion of the gradient in the transfer layers between neurons, in which the iteration control, smaller error and better model are stopping criteria of the used algorithm. The functions used are shown below:

• Linear;

• Sigmoid;

• Hyperbolic tangent

3.2 Methods

The research is characterized by the investigation of computational methods that show efficiency in data processing and effectiveness in simulating the results, to apply the concepts related to the forecast of the river level in function of the scenarios of environmental impacts caused by floods or ebbs, through database referring to the river level in the state of Amazonas, an analysis was carried out on the content.

3.2.1 Survey of Data Relating to the River Level

The data collection must be done in meteorological databases, or specific sites that show the collection of data in a chronological way, in such a way that it is possible to use this recorded history.

The mapping of the input variable must be done through the National Institute of Meteorology by consulting the database, the output variable must be acquired through the Port of Manaus website.

3.2.2 Selection of Computational Methods for the Forecasting Model

In this step, the computational methods that should be used for the NARX learning process are defined, in this case, specifically because they are time series with seasonality, and stochastic variables related to the river level, it is necessary to use the model with inputs exogenous autoregressive.

The algorithms used to test the feasibility of using NARX were: Optimization, Heuristic, Iterative, Quasi-Newton, Newton and Numerical, these methods are of paramount importance for performing mathematical, stochastic procedures and iterative searches to find local and close to global solutions. , targeting minimization or maximization factors with respect to algorithm performance on a dataset.

3.2.3 Definition and Implementation of the NARX Architecture

As shown in Figure 3, the NARX architecture is composed of three (3) input variables, namely: Rainfall, Temperature and Relative Humidity, in addition, one (1) output variable, which is the river level. The concept of feedback, or Recurrent Neural Network, is due to the fact that the architecture provides

mechanisms to recursively call a function or pass the estimated parameter to a new training cycle without losing the a priori adjustment information.

Figure 3 - NARX Architecture.
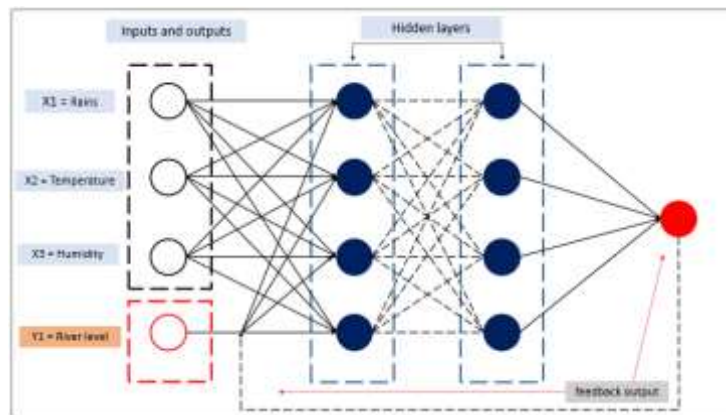
Source: (AUTHOR, 2021).

A regressive model, corresponds to the NARX architecture, through state space models, the transition function is responsible for mapping the current state and the current input of the dynamic system, in such a way as to return this without loss of information, characterizing a state transition (MUÑOZ CHÁVEZ, 2020), it can be represented by Eq. (3.1) and (3.2).

$$x(n + 1) = Ax(n) + Bu(n) \tag{3.1}$$

$$y(n0 = Cx(n) \tag{3.2}$$

In this way, a regressive model like NARX has its terminologies similar to a statistical model, in such a way that the computational model reaches an expected and estimated degree of precision, thanks to the implemented mathematical models that give support and consistency in the approximation of functions to the use a training algorithm. Table 3 presents a comparison of the terminologies of a NARX regression model and a statistical model.

Table 3 - Terminology between an ANN and STATISTICAL model.

NEURAL NETWORK	STATISTIC
Appetizer	Exogenous Variables
outputs	Endogenous Variables
weights	parameters
training set	Sample
backpropagation	stochastic approximation
Training	pet

Source: (VILELA and MATEUS, 2016).

3.2.4 Simulation and Comparison of Results with the Winning Algorithm

As shown in Figure 4, for the execution of the simulation procedures and comparison of the results with the winning algorithm, 3 steps will be necessary, namely: Selection of n layers and best transfer function; Selection of n neurons, number of feed delays and use of Parallelism; finally, Simulation with 12 training algorithms.

Figure 4 - Steps for simulation and comparison of results.
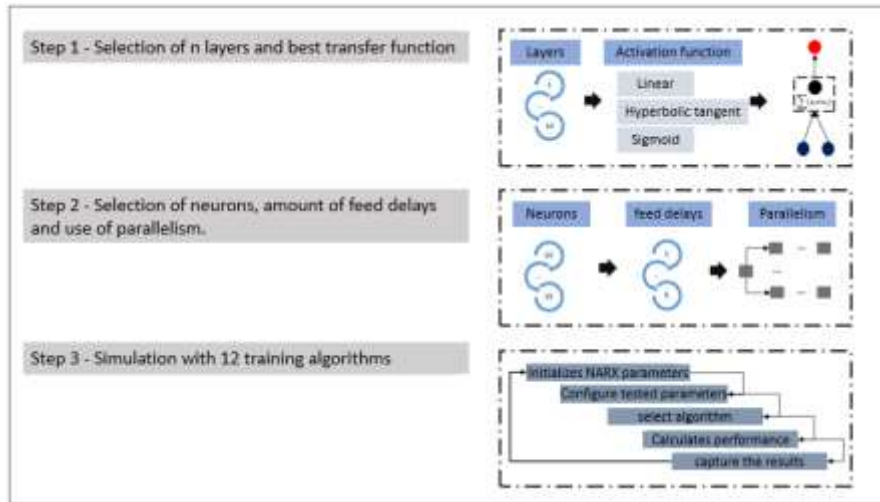
Source: (AUTHOR, 2021).

The first step consists of performing verification and testing procedures to find the ideal number of layers and the best transfer function, based on a simulation range from 1 to 10, where Linear, Hyperbolic Tangent and Sigmoid transfer functions are used.

With this, it is expected to obtain satisfactory results due to a number of 30 Neural Networks trained and able to perform predictions, in this way the developed algorithm must analyze and point out the best ANN architecture to be used in the next step.

The second step consists of selecting an amount x of neurons, an amount y of feed delays and the use of parallelism or not. Due to the demand as a function of time and computational cost, it is necessary to take advantage of the auxiliary processors of the test machine, characterizing parallel computing.

In this way, it is expected to find the satisfactory amount in a range of 10 to 15 neurons, a satisfactory amount of feed delays ranging from 1 to 5 and the definition of the computational cost for using parallel computing.

Finally, the third step consists of actually training the best ANN architecture using a method of statistical analysis by calculating 12 training algorithms, in such a way that it is possible to define the best ANN with the optimized algorithm depending on the quantity characteristics. of layers, neurons, activation functions and use of parallelism.

The algorithm will have as main workflow, 5 steps that lead to the generation of results, that is, graphs, tables and performance indicators according to training and simulations.

Step 1 – Initializes the NARX parameters;

Step 2 – Configure the tested parameters;

Step 3 – Select the current algorithm n of 12;

Step 4 – Calculates and analyzes training performance;

Step 5 – Capture the results.

To statistically analyze the ANN training performance as a function of its characteristics, in this case, calculated model and approximate model, it is done through four mathematical models to measure errors through the waste flow and dispersion rate, called:

• Regression;

• Mean Square Error (MSE);

• Root Mean Square Error (RMSE);

• Mean Absolute Percentage Error (MAPE).

Expressed respectively by Eqs. (3.3), (3.4) and (3.5).

$$MSE = \frac{\sum_{i=1}^{n}(ra_i - rs_i)^2}{n} \tag{3.3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(ra_i - rs_i)^2}{n}} \tag{3.4}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{ra_i - rs_i}{ra_i}\right| \tag{3.5}$$

Where:

MSE = Mean Square Error;

RMSE = Root Mean Square error;

MAPE = Mean Absolute Percent Error;

n = number of elements in the output vector;

ra = target result;

rs = simulated result;

Y = is the dependent variable and $\hat{Y}\_t$ is the predicted value of Y given an X_i;

X = is the independent variable;

α = is the predicted value of Y when X = 0 (the intercept);

β = is how much Y changes, on average, per unit change in X (the slope).

Due to the use of statistical models to analyze the error rates, a mathematical model expressed by Eq. (3.6) was prepared in order to facilitate the process of selecting the best ANN architecture based on these

indicators, in such a way that, the lower the value acquired by this indicator, the better the performance.

$$Performance = (1 - Regressão) + MSE + RMSE + MAPE \qquad (3.6)$$

In the end, the algorithm should generate analysis results with performance indicators of the simulation performance and of the scenarios projected for the river level prediction problem.

4. RESULTS AND DISCUSSIONS.

The partial results are in accordance with the steps planned for the research, where it was carried out in order to validate the proposal and the hypothesis of a solution at work. For the qualification project, it was possible to acquire:

    • The database containing the meteorological records necessary for acquiring the forecast model;

    • The significant variables that will serve as the input history of the NARX model for the river level forecast;

    • The definition of the algorithm responsible for the river level prediction, computationally characterized as a prediction model.

4.1 Database

The database is composed of 4 variables, namely: Rain, Temperature, Humidity and River Level, Table 1 presents a sample of this database. It is worth noting that this table presents a stratification of data from 01/01/2000 to 06/30/2021 collected from the bases of the Port of Manaus and INMET.

Table 1 - Data sample.

Date	Rain	Temperature	Humidity	River level
31/01/2000	350,2	25,87	83,93	21,22
29/02/2000	344,4	25,69	84,58	22,46
31/03/2000	340,7	25,70	86,14	24,4
30/04/2000	535,4	25,51	86,97	26,56
31/05/2000	172,6	26,33	85,31	28,05
31/10/2020	285,8	33,4	83,6	17,16
30/11/2020	197,8	33,9	86,2	16,9
31/12/2020	260,6	33,1	89,7	19,3
31/01/2021	432,6	26,3	89,3	22,98
28/02/2021	583,2	26,6	83,6	25,19
31/03/2021	588,9	27	84,6	26,45
30/04/2021	426,8	26,4	88,2	28,15
31/05/2021	488,8	26,5	87,8	29,63
30/06/2021	89,1	27,2	79,2	30,02

Source: Adapted from (PORTO DE MANAUS, 2021; INMET, 2021).

4.2 Definition of the algorithm

Figure 5 shows the flowchart of steps of the main algorithm, with five steps for acquiring the final results.

Figure 5 - Algorithm flowchart.
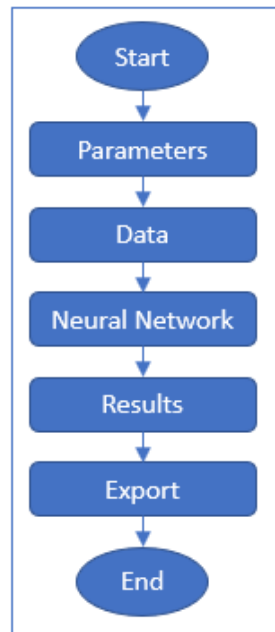
Source: (AUTHOR, 2021).

Step 1 - Definition of parameters according to the characteristics of the NARX model;

Step 2 – Reading the database and selecting significant variables for the forecast model, for training and simulation;

Step 3 – Training of the Recurrent Neural Network, using the NARX model, definitions of the configurations of the training algorithms, data division and performance analysis;

Step 4 - Execution of simulations according to the best Neural Network architecture and generation of results;

Step 5 - Export of graphs, tables and indicators of simulation results and forecast model performance.

4.2.1 Algorithm Performance Tests

The first test aimed to find the ideal number of layers for the computational model and the transfer function with the best performance rate. Thus, the Levenberg-Marquardt training algorithm was used, where an interval of layers between 1 to 10 is tested, alternating between the 3 transfer functions: Linear, Hyperbolic Tangent and Sigmoid, Figure 6 shows the test scheme.

Figure 6 - Test scheme for layers.
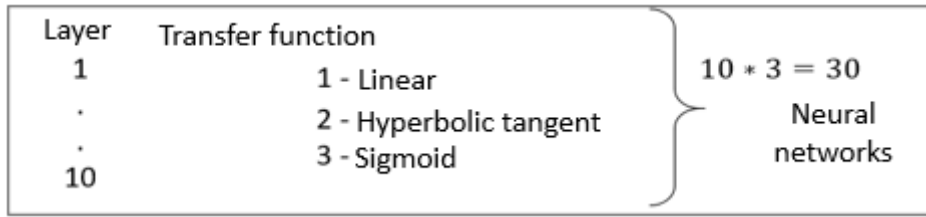
Source: (AUTHOR, 2021).

Through the scheme, the results found are shown by Figures 7, 8 and 9, respectively with the functions: Linear, Sigmoid and Hyperbolic Tangent. The graph in Figure 7 shows expressive values when increasing the number of layers, as a result, the performance rate follows a varying pattern between 2 and 3, however, the time in seconds is also greater as the number of layers is increased. .

Figure 7 - Linear function graph for the Layers test.
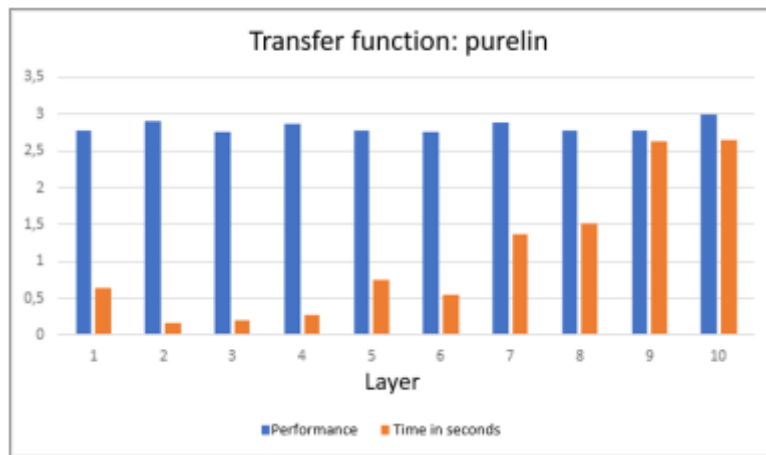
Source: (AUTHOR, 2021).

Figure 7 shows the graph of the Sigmoid function, it presents a different behavior compared to Figure 8, where the time in seconds increases as a function of the number of layers, however, the performance rate remains variable in a range of 6 to 7, it is possible to identify that the number of layers 7 presents the lowest results characterizing the best in relation to the range 1 to 10.

Figure 8 - Sigmoid function graph for the Layers test.
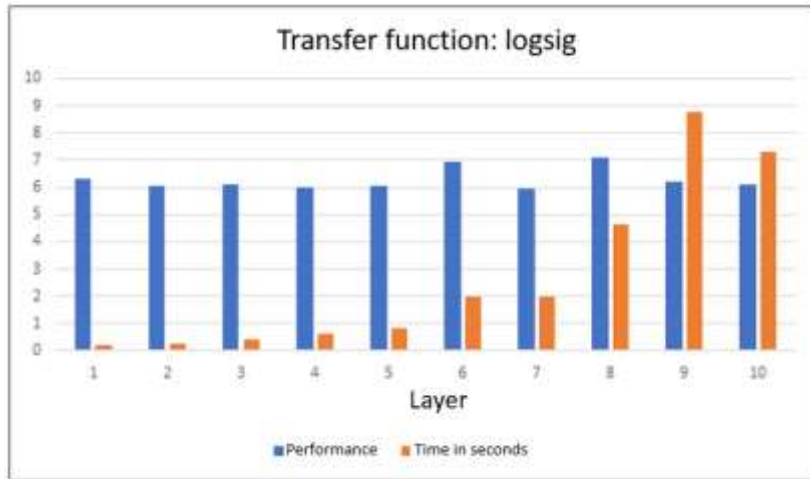
Source: (AUTHOR, 2021).

Figure 9 shows the graph of the Hyperbolic Tangent function. other figures, where the variance occurs in a range of 1.5 to 4.3, it is possible to identify that even layer 5 presents the best results for performance and time in seconds.

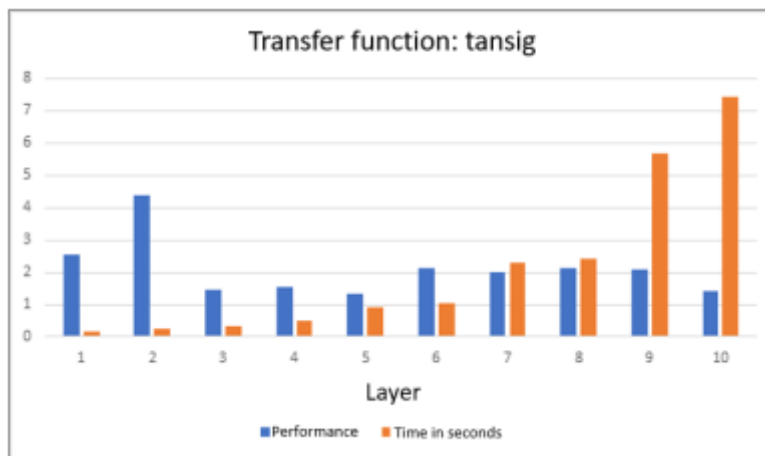Figure 9 - Graph of the Hyperbolic Tangent function for the Layers test.

Table 2 presents the values of the performance simulation results performed for the 3 transfer functions using the Levenberg-Marquardt learning algorithm, where it is possible to identify that the best (winner) is the Layer 5 Hyperbolic Tangent, with a time of 0, 90328, 17 epochs and a rate of 1.33307 for performance.

Table 2 - Results of the performance simulation.

Nº	Layers	Trans. Func.	Time	Epochs	Perform ance	Reg.	MSE	RMSE	MAPE
1	**5**	**tansig**	**0,90328**	**17**	**1,33307**	**0,98243**	**0,55023**	**0,74177**	**0,02349**
2	10	tansig	7,42808	21	1,42664	0,98051	0,60514	0,77791	0,02411
3	3	tansig	0,33391	13	1,46888	0,98027	0,62997	0,79371	0,02546
4	4	tansig	0,48584	13	1,53483	0,97797	0,66811	0,81738	0,02731
5	7	tansig	2,29151	22	1,98561	0,96884	0,94815	0,97373	0,03257

The second test aimed to identify the ideal number of neurons for the layers, in this case 5 layers, having seen the result of the first test, in addition, the use of parallel computing is verified and the number of feed delays, as a result, where 1765 neural network models were totaled as a test for this step, as shown in Figure 10.

Figure 10 - Test scheme for neurons.
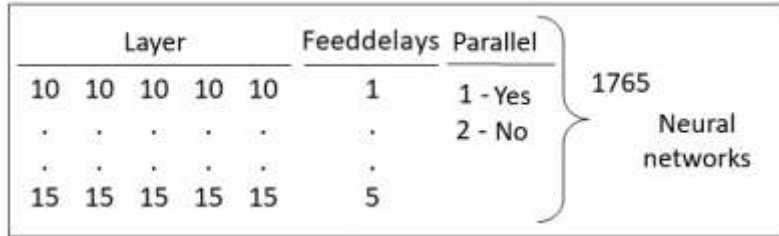
Source: (AUTHOR, 2021).

Table 3 - Neuron simulation results.

Nº	Feedback Delays	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Epochs	Parallel	Fun. Perf.	Tempo (s)
1	**2**	**15**	**10**	**15**	**10**	**10**	**27**	**F**	**0,9181**	**1,6712**
2	3	12	12	13	13	13	23	F	0,9724	1,7684
3	5	15	15	15	10	15	20	F	1,0565	4,0492
4	2	13	12	13	13	12	25	V	1,0690	6,6479
5	3	11	11	14	14	11	25	F	1,0887	1,9040

Source: (AUTHOR, 2021).

Table 3 presents the results of the test to find the winning parameters for the best number of neurons that will be used with 5 layers in the computational model, it is possible to identify that the model with 2 feed delays, 15 neurons in the first layer, 10 in the second, 15 in the third, 10 in the fourth and 10 in the fifth, a performance rating of 0.9181 and a time of 1.6712 seconds, was the winner.

5. Conclusion

Through the results found with the computer simulations to find the ideal number of layers and the number of neurons, it is possible to affirm that the simulations were important to verify the performance and quality indicators of the computational model in function of the chosen training algorithm (Levenberg -Marquardt). On the other hand, of the four specific objectives it was possible to develop two of them, in such a way that the significant variables for the forecast model were acquired, as mentioned in the materials section of chapter 3, the selection of indicator settings for the computational model (quantity number of layers, number of neurons, activation functions and number of feed delays) were acquired.

Finally, the results presented here are of relevant characteristics for the continuity of the research and improvement of the parameters with other training algorithms to carry out comparisons of results with the NARX and SARIMA model in order to achieve the general objective of applying this computational model

to the prediction of the river level due to the environmental impacts caused by floods or ebbs in the state of Amazonas.

6. Acknowledgement

7. References

ALBERTON, Gabriel Barreto et al. Aplicação de redes neurais artificiais para previsão de enchentes no rio Itajaí-Açu em Blumenau, SC, BrasiI. **Revista Ibero-Americana de Ciências Ambientais**, v. 12, n. 4, 2021.

ALBUQUERQUE, Francisco Roberto. Os problemas causados pela cheia do rio Amazonas na área do bairro da Francesa na cidade de Parintins no ano de 2015. 2018.

ALENCAR, David B. et al. Hybrid approach combining SARIMA and neural networks for multi-step ahead wind speed forecasting in Brazil. **IEEE Access**, v. 6, p. 55986-55994, 2018.

ALVES, Luna Gripp Simões. Relatório para estabelecimento de cotas de referência para alerta hidrológico em Municipios da Amazônia Ocidental. 2021.

ARAÚJO, Carla Beatriz Costa de et al. Previsão Sazonal de Vazões para a Bacia do Orós (Ceará, Brasil) Utilizando Redes Neurais e a Técnica De Reamostragem dos K-vizinhos. **Revista Brasileira de Meteorologia**, v. 35, p. 197-207, 2020.

BARBOSA DE ALENCAR, David et al. Different models for forecasting wind power generation: Case study. **Energies**, v. 10, n. 12, p. 1976, 2017.

BLÁZQUEZ-GARCÍA, Ane *et al.* Short-term office building elevator energy consumption forecast using SARIMA. **Journal of Building Performance Simulation**, v. 13, n. 1, p. 69-78, 2020.

CERRI, Ricardo. Introdução às Redes Neurais Artificiais com Implementações em R. In: **Anais da I Escola Regional de Aprendizado de Máquina e Inteligência Artificial de São Paulo**. SBC, 2020. p. 47-50.

CHANG, Wen-Yeau et al. A literature review of wind forecasting methods. **Journal of Power and Energy Engineering**, v. 2, n. 04, p. 161, 2014.

CRISTALDO, Marcia Ferreira et al. [artigo retratado] redes neurais artificiais aplicadas à previsão de enchentes para região do pantanal no mato grosso do sul. **Geosciences**= **Geociências**, v. 39, n. 1, p. 191-201, 2020.

DE AMORIM, Renata Ferreira. A gestão dos recursos hídricos através da tríade soberania, sustentabilidade e desenvolvimento sustentável–recorte na Região Hidrográfica Amazônica. **Anais dos Seminários de Iniciação Científica**, n. 23, 2021.

DE MENDONÇA, Leonardo Melo et al. Modelagem chuva-vazão via redes neurais artificiais para simulação de vazões de uma bacia hidrográfica da Amazônia. **Revista de Gestão de Água da América Latina**, v. 18, n. 2021, 2021.

FAVA, Maria Clara. **Modelo de alerta hidrológico com base participativa usando sistema de informações voluntárias para previsão de enchentes**. 2015. Tese de Doutorado. Universidade de São

Paulo.

FINCK, Juliano Santos. Previsão em tempo atual de níveis fluviais com redes neurais artificiais: aplicação à bacia do Rio Taquari-Antas/RS. 2020.

GORODETSKAYA, Yulia; DA FONSECA, Leonardo Goliatt; DE MELO RIBEIRO, Celso Bandeira. PREVISÃO DE VAZÃO DE CURTO PRAZO UTILIZANDO REDES NEURAIS ARTIFICIAIS. ESTUDO DE CASO: RIO PARAÍBA DO SUL, 2017.

HAYKIN, S. O. Neural Networks and Learning Machines. 3 Prentice Hall. **New York**, 2008.

INMET. 2021. Acessado em: 18/08/2021. Disponível em: https://bdmep.inmet.gov.br/

MARACAJÁ, José Rosenilton de Araújo et al. Previsão regionalizada de vazão sazonal utilizando redes neurais artificiais. 2005.

MATKOVSKYY, romano; BOURAOUI, Taoufik. Aplicação de redes neurais a índices compostos de séries temporais curtas: Evidência do modelo não linear autoregressivo com entradas exógenas (narx). **Journal of Quantitative Economics**, v. 17, n. 2, pág. 433-446, 2019.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, v. 5, n. 4, p. 115-133, 1943.

MUÑOZ CHÁVEZ, Jairo José. Modelagem em espaço de estados e predição da geometria do cordão de solda no processo GMAW usando redes SVM, neurodifusa e recorrentes de alta ordem. 2020.

NATTRODT, Thaísy Nitis Mota; DIAS, Maria Das Graças Santos. As relações entre os recursos hídricos, a energia e a sustentabilidade na Amazônia. **Brazilian Journal of Development**, v. 7, n. 4, p. 38319-38339, 2021.

PARENTE, Ricardo Silva et al. Application of the narx model for forecasting wind speed for wind energy generation". **International Journal of Development Research**, v. 11, n. 04, p. 46461-46466, 2021.

PORTO DE MANAUS, O CORAÇÃO DO AMAZONAS. 2021. Acessado em: 18/08/2021. Disponível em: https://www.portodemanaus.com.br/?pagina=nivel-do-rio-negro-hoje

REIS, Leonardo Pequeno et al. Prognose da distribuição diamétrica na Amazônia utilizando redes neurais artificiais e autômatos celulares. **Floresta**, v. 48, n. 1, p. 93-102, 2018.

Serviço Geológico do Brasil (CPRM). **Boletim de Monitoramento Hidrometereológico da Amazônia Ocidental**. 2021.

SILVA, João Batista Lopes da. Modelos de previsão de enchentes em tempo real para o município de Nova Era MG. 2006.

SILVA, Marcos Cardoso et al. Aplicação de redes neurais artificiais na cultura do mogno (Khaya spp. e Swietenia spp.). **Journal of Biotechnology and Biodiversity**, v. 8, n. 1, p. 017-023, 2020.

SILVA, Maria do Socorro Rocha da et al. Bacia hidrográfica do Rio Amazonas: contribuição para o enquadramento e preservação. 2013.

SILVA, Maria do Socorro Rocha da; MIRANDA, Sebastião Átila Fonseca; SANTANA, Genilson Pereira. Bacia Hidrográfica do Rio Amazonas: Condições de suas águas versos Resolução n° 357/CONAMA/2005. **Volume 6, Pags**. 83-90, 2016.

VILELA, Leticia Biagi; MATEUS, Tiago Henrique de Abreu. APLICAÇÃO DE REDE NEURAL NARX PARA A PREVISÃO DO PREÇO DA SOJA. **XXXVI Encontro nacional de engenharia de produção**. 2016.

WANG, Wen-chuan *et al.* Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. **Water Resources Management**, v. 29, n. 8, p. 2655-2675, 2015.

WUNSCH, Andreas; LIESCH, Tanja; BRODA, Stefan. Previsão dos níveis de água subterrânea usando redes autorregressivas não lineares com entrada exógena (NARX). **Journal of hydrology**, v. 567, p. 743-758, 2018.

Copyright Disclaimer