Research Article

Improved Rainfall Prediction through Nonlinear Autoregressive Network with Exogenous Variables: A Case Study in Andes High Mountain Region

Mario Peña,^{1,2} Angel Vázquez-Patiño,^{3,4} Darío Zhiña,⁵ Martin Montenegro,⁵ and Alex Avilés⁵

¹Dirección de Investigación (DIUC), Universidad de Cuenca, Campus Central, Av. 12 de Abril s/n y Loja 010203, Cuenca, Ecuador

²Departamento de Química Aplicada y Sistemas de Producción, Facultad de Ciencias Químicas, Universidad de Cuenca, Campus Central, Av. 12 de Abril S/n y Loja 010203, Cuenca, Ecuador

³Facultad de Ingeniería, Universidad de Cuenca, Av. 12 de Abril S/n y Loja 010203, Cuenca, Ecuador

⁴Departamento de Ingeniería Civil, Universidad de Cuenca, Av. 12 de Abril S/n y Loja 010203, Cuenca, Ecuador

⁵Carrera de Ingeniería Ambiental, Facultad de Ciencias Químicas, Universidad de Cuenca, Campus Central, Av. 12 de Abril S/n y Loja 010203, Cuenca, Ecuador

Correspondence should be addressed to Alex Avilés; alex.aviles@ucuenca.edu.ec

Received 3 December 2019; Revised 4 July 2020; Accepted 1 September 2020; Published 17 September 2020

Academic Editor: Roberto Coscarelli

Copyright © 2020 Mario Peña et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Precipitation is the most relevant element in the hydrological cycle and vital for the biosphere. However, when extreme precipitation events occur, the consequences could be devastating for humans (droughts or floods). An accurate prediction of precipitation helps decision-makers to develop adequate mitigation plans. In this study, linear and nonlinear models with lagged predictors and the implementation of a nonlinear autoregressive model with exogenous variables (NARX) network were used to predict monthly rainfall in Labrado and Chirimachay meteorological stations. To define a suitable model, ridge regression, lasso, random forest (RF), support vector machine (SVM), and NARX network were used. Although the results were “unsatisfactory” with the linear models, the specific direct influences of variables such as Niño 1 + 2, Sahel rainfall, hurricane activity, North Pacific Oscillation, and the same delayed rainfall signal were identified. RF and SVM also demonstrated poor performance. However, RF had a better fit than linear models, and SVM has a better fit than RF models. Instead, the NARX model was trained using several architectures to identify an optimal one for the best prediction twelve months ahead. As an overall evaluation, the NARX model showed “good” results for Labrado and “satisfactory” results for Chirimachay. The predictions yielded by NARX models, for the first six months ahead, were entirely accurate. This study highlighted the strengths of NARX networks in the prediction of chaotic and nonlinear signals such as rainfall in regions that obey complex processes. The results would serve to make short-term plans and give support to decision-makers in the management of water resources.

1. Introduction

Precipitation is one of the main components in the hydrological cycle [1, 2]. It is one of the most important variables associated with atmospheric circulation in meteorological studies [3]. Moreover, it is the main source of recharge in water balance studies from local to regional scales [4]. However, its forecast has become a significant

challenge owing to the complexity of the atmospheric processes that produce rainfall [5]. This reality is evident in high mountain regions [6] with high time-space variability like the Andes mountain range. In particular, the water originating from the Andean high mountains is used for diverse purposes (e.g., human consumption, agriculture, livestock grazing, industry, and recreation) in downstream cities [7–9] (e.g., Mérida in Venezuela, Bogotá in Colombia,

and Quito in Ecuador). An accurate prediction of precipitation (temporal and spatial) can help decision-makers to assess in advance both flood and drought situations [10, 11], and it could support extreme hydrological management and diminish the impacts on the population.

In different studies, a variety of methods have been used to forecast rainfall. Some of them have shown accurate results like the multiple linear regression and k -nearest-neighbors methods as in [12, 13]. Other approaches have used the outputs of global climate models to improve the seasonal and subseasonal forecasting [14–16] and the probabilistic forecasts for uncertainty quantification [17–19]. Other studies have applied artificial intelligence approach for rainfall prediction such as artificial neural networks (ANNs) [5, 10–12, 20–24], support vector machine (SVM) [12, 13, 24, 25], logistic regression [26], and random forest [27, 28]. Also, other studies have investigated the optimum selection of predictors to improve forecast accuracy [29]. NARX model is a novel approach for many applications of prediction of environmental variables, for example, water level of wetland systems [30], groundwater levels [31–33], urban drainage framework [34], stream flow and inflow to reservoirs [35, 36], water level in urban flood [37–39], and rainfall forecasting [40, 41].

The unpredictable and chaotic nature of the rainfall behavior makes its prediction a critical challenge. The NARX model is suitable for dealing with time series with such characteristics [42] since the output value of a variable is a nonlinear function of the past values of the same variable and other exogenous variables. Nonlinear mapping is convenient in chaotic environments, and it is superior to conventional autoregressive models. It is also superior to ANN since NARX uses its memory to make predictions. NARX presents an improvement opportunity due to its performance in the prediction of time series that has a seasonal component.

In Andean countries like Ecuador, some of the methods mentioned above have also been used to predict precipitation. For example, in [43], continuous transfer function models were used to predict rainfall on the coasts of Ecuador. Also, Mendoza et al. [44] used sea surface temperature (SST) as a predictor and canonical correlation as a method to predict rainfall in the coast and Andes of Ecuador. These studies highlight the complex relationship between predictors and rainfall in the Andes. However, new methods like NARX could be applied in the Ecuadorian Andes to improve the performance of precipitation forecasting.

This study first explores the performance of the most commonly used linear models for the rainfall forecast. It then compares them with nonlinear models to verify their performance in a mountain basin located in the south of Ecuador. Finally, NARX models are used to forecast rainfall several steps ahead. The results obtained can have a significant impact on mountain systems because they could be used to obtain improved performance in high mountain areas.

2. Materials and Methods

2.1. Study Area. The study area is composed of the high mountain subbasins Tomebamba and Machángara, located in the Andean southern Ecuador (Figure 1). The two subbasins are essential to supply water to the city of Cuenca, considered the third most important in Ecuador [45]. Moreover, the two subbasins belong to the Paute hydrographic system, which, together with others, provide water to generate energy. The subbasins altitude varies between 2440 and 4400 m.a.s.l., which in turn allows finding a great variety of vegetation. For example, in high areas, there are patches of *Polylepis reticulata*, tussock grasses, ground rosettes, cushion plants, and ground rosettes, while in the low areas, there are agricultural land, pastures, as well as urban zones [9, 46, 47].

Two meteorological stations are taken into account for the analysis: Labrado and Chirimachay. These are located in the upper parts of the two subbasins at 3300 m.a.s.l (Figure 1). According to Celleri et al. [48], there are two types of precipitation within the study area: bimodal type I in the lower parts of the subbasins and bimodal type II in the middle and upper zones. Both regimes have two peaks of precipitation during April and October; however, the bimodal type II regimen presents a less dry season from June to August (Figure 2).

2.2. Precipitation Data. Rainfall data from Labrado (subbasin of Machángara) and Chirimachay (subbasin of Tomebamba) stations were provided by the Ecuadorian National Institute of Meteorology and Hydrology (INAMHI, <http://www.serviciometeorologico.gob.ec>). The study period is 1964–2015. Table 1 shows a statistical summary of the rainfall data.

2.3. Climatic Indices. The 27 climatic indices listed next were used in the study: The Atlantic Meridional Mode (AMM) SST Index [49]; Atlantic Multidecadal Oscillation (AMO) [50]; Antarctic Oscillation (AO) [51]; Bivariate ENSO Time series (BEST) [52]; Caribbean Index (CAR) [53]; hurricane activity (HURR) [54]; Multivariate ENSO Index (MEI) [55]; North Atlantic Oscillation (NAO) [56]; Niño 1 + 2, Niño 3, Niño 3.4, and Niño 4 [57]; North Pacific Oscillation (NP) [58]; North Tropical Atlantic SST Index (NTA) [53]; Oceanic Niño Index (ONI) [59]; Pacific Decadal Oscillation (PDO) [60]; Pacific North American Index (PNA) [61]; Quasibiennial Oscillation (QBO) [62]; Sahel rainfall (SAHELRAIN) [63]; Southern Oscillation Index (SOI) [64]; solar flux [65]; Tropical Northern Atlantic Index (TNA) [66]; Trans-Niño Index (TNI) [67]; Tropical Southern Atlantic Index (TSA) [68]; western hemisphere warm pool (WHWP) [69]; Western Pacific Index (WP) [70]; and global mean land/ocean temperature (GMSST) [71, 72]. Time series of these indices are available in <https://www.esrl.noaa.gov/psd/data/climateindices/list/>.

FIGURE 1: Location of the study area made up of the subbasins Machángara and Tomebamba.

FIGURE 2: Seasonality of rainfall in (a) Labrado and (b) Chirimachay stations.

TABLE 1: Statistical summary of stations rainfall data.

Station	Standard deviation	Maximum	Minimum	Mean	Skewness	Kurtosis
Labrado	45.65	287.30	12.80	104.06	0.68	3.46
Chirimachay	57.33	490.30	1.90	111.90	1.79	9.75

2.4. Linear Models

2.4.1. Linear Model. The linear model predicts a quantitative output y based on a single predictor x , assuming that a linear relationship between them exists. The following equation describes this linear relationship:

$$y = \beta_0 + \beta_1 x + \varepsilon, \quad (1)$$

where β_0 is the bias (offset), β_1 is the coefficient (slope) of variable x , and ε is the error or random noise.

2.4.2. Multivariate Linear Model. In general, the multivariate linear model supposes that p distinct predictors are available, and the output is a weighted linear combination of the set X of predictor variables x . The formula of the multivariate linear model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon, \quad (2)$$

where x_j represents the j th predictor and β_j represents the magnitude of the relationship between the j th predictor and the output y . The absolute value of the coefficients β_j defines the degree of influence of the predictor over the output [73].

2.4.3. Linear Model Regularization and Selection. As the availability of variables increases, the probability of falling into overfitting also increases [74]. Overfitting is an error that occurs when a model fits too closely to a limited set of data points, decreasing the predictive power of the model. To prevent this drawback, commonly, ridge and lasso regressions are used as regularization techniques [73]. The objective of applying these techniques is to obtain parsimonious models. In the multivariate linear model, the adjustment of the parameters is made by minimizing the cost function (residual sum square, RSS) through the least-squares. The RSS formula is shown as follows:

$$\text{RSS} = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2, \quad (3)$$

where n is the number of samples in the dataset and x_{ij} is the value of the i th sample in the j th predictor. Ridge regression aggregates a penalty term to the cost function, which is equal to the sum of squares of the magnitude of the coefficients. As a result, this method keeps all the predictors with the lowest coefficient magnitudes. In the lasso regression approach, the penalty term is equal to the sum of the magnitude of the absolute coefficient, and some values even can become zero. This is the reason why Lasso is considered as a feature

selection method. Ridge regression and lasso minimize the quantity depicted on the following equations, respectively:

$$\text{RSS} + \sum_{j=1}^p \beta_j^2, \quad (4)$$

$$\text{RSS} + \lambda \sum_{j=1}^p |\beta_j|, \quad (5)$$

where $\lambda \geq 0$ is a tuning parameter. The impact of regularization over the estimated coefficient is controlled through λ that accompanies the penalty term. In an extreme case, when $\lambda = 0$, the penalty term does not cause any effect. Conversely, as $\lambda \rightarrow \infty$, the incidence of the regularization penalty increases [73].

2.4.4. Training and Testing Datasets. The data of the 27 synoptic climatic indices were lagged 12 months with respect to rainfall data. This decision is due to a practical matter. To predict 12 months of rain, information on exogenous variables from the previous 24 to 12 months is necessary. From that, the information was lagged from 1 to 24 months, so there were 675 time series used as predictors corresponding to the climatic indices. To incorporate the sense of autoregression in linear models, the same rainfall signal was used as a predictor in the models that considered lagged exogenous predictors. Although some predictors were correlated with each other (e.g., Spearman index of 0.94 between TNA and NTA), none were omitted. This is because only perfectly correlated signals (i.e., genuinely redundant) do not provide additional information [75]. The original dataset was divided into two subsets: the first one from January 1964 to December 2014 and the second one from January 2015 to December 2015. The Minmax normalization process was applied to the first subset, which produced parameters to normalize the second subset as well. The first subset was randomly divided into a subset to train models (80% to find the best coefficients) and a subset to test (20% to assess the model). Applying this approach, fifty multivariate linear models to estimate the rainfall were fitted, namely, twenty-five for each method (ridge and lasso), one with no lags, and 24 with lags ranging from 1 to 24 lags. The algorithm used to define the best value of λ and fit the models was cross-validation [73].

2.5. Random Forest. Random forest (RF) [76] is based on the idea of boosting the prediction of a model using an assembly of decision trees [77] results. Each random tree is based on an independent random vector of sample values with the same distribution. RF has become popular in hydrological and climatic applications (e.g., [27, 28]) due to its high performance, efficient training in big datasets and high

dimensions, and the estimation of the importance of the different features that represent the instances.

Several hyperparameters can be optimized in the construction of the model based on random forest. From them, different values for the number of decision trees, its maximum depth, and the number of features used in the construction of the trees were explored in a grid search 6-fold cross-validation [78] fashion. For this, 85% of data were used as a training subset. For testing the performance of the resulting model, the remaining independent 15% was used. Finally, the model was used for the prediction of rainfall in 2015 in the two study stations.

2.6. Support Vector Machines. Support vector machine (SVM) [79] raises the dimensionality of the data to a vector space where it is possible to construct a linear regression model. The linear regression is performed based on representative data points that make up a so-called support vector. The representation of the data in a high dimension is done using kernel functions. Here, the Gaussian radial basis function (RBF) was used. Two hyperparameters must be optimized for the regularization (parameter C) and the spread of influence of the support vector (parameter gamma).

The optimization of the hyperparameters was done in the same way as with RF, i.e., 85% of data for training and 6-fold cross-validation and 15% for testing. The resulting model was used for predicting the rainfall in 2015 in the two study stations. Many studies in climate and meteorology have used SVM [12, 13, 24, 25] and RF in prediction, so we use them here as a baseline to compare the performance of NARX.

2.7. Recurrent Neural Nets with Exogenous Inputs NARX Model. The nonlinear autoregressive network with exogenous inputs (NARX) is a dynamic recurrent neural network (RNN), with feedback connections that enclose several layers of the network; i.e., the output is considered as another input of the network. Figure 3 depicts the NARX model architecture.

Its memory ability is useful for the prediction of nonlinear time series. Besides, unlike classic artificial neural networks, NARX gains degrees of freedom by incorporating valuable information from exogenous inputs [32]. There are two different architectures of NARX: the series-parallel architecture (called open-loop) and the parallel architecture (called close-loop) presented by the following equations, respectively:

$$\begin{aligned} \hat{y}(t+1) = F(y(t), y(t-1), \dots, y(t-n_y), \\ X(t+1), X(t), \dots, X(t-n_x)), \end{aligned} \quad (6)$$

$$\begin{aligned} \hat{y}(t+1) = F(\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t-n_y), \\ X(t+1), X(t), \dots, X(t-n_x)), \end{aligned} \quad (7)$$

where $F(\cdot)$ is the mapping function and $\hat{y}(t+1)$ is the predicted output of NARX for the time $t+1$, determined at the instant t . The terms $y(t)$, $y(t-1)$, $\dots$, $y(t-n_y)$ are the

true past values of the time series, also known as the ground truth, and $\hat{y}(t)$, $\hat{y}(t-1)$, $\dots$, $\hat{y}(t-n_y)$ is the past predicted outputs generated by the NARX. The true values of exogenous inputs are $X(t+1)$, $X(t)$, $\dots$, $X(t-n_x)$. The numbers of input delays and output delays are defined by n_x and n_y , respectively.

The series-parallel architecture is used in the training process because the real output is available, leading to a conventional feed-forward representation, while the parallel architecture can make predictions based on feeding back the estimated output instead of accurate output.

The mapping function $F(\cdot)$ (which is initially unknown) is fitted as the training process progresses. The multilayer perceptron (MLP) architecture is used to represent this approximation since it is a robust structure capable of learning any kind of continuous nonlinear mapping. A classic MLP contains three basic layers: input, hidden, and output layers. Moreover, it has elements such as neurons, activation functions, and connections weights. As the number of hidden neurons increases, the model can approach more complex functions. However, the selection of the number of these hidden neurons depends on the addressed case study. In general, just one hidden layer with neurons ranging from $[0.5p, 2p]$ is commonly used [80]. The sigmoid-linear transfer function combination can provide an efficient mathematical representation of the output as a function of the input signal [32].

2.7.1. NARX Model Architecture. In NARX models, 27 synoptic climate indices and the rainfall signal were used as inputs to the nets, and rainfall was the output. Again, none of the predictor was excluded. To identify the optimum NARX architecture and due to the massive amount of net setting allowed, the standard trial-and-error method to select the number of hidden nodes and lags number was used. The input layer neurons depended on the number of lags used, so architectures with 2, 3, 4, 5, 6, 9, and 12 lags were tested. The output layer had just one neuron. The net had one hidden layer with 10, 20, 30, 40, and 50 neurons. The neurons in the hidden layer used a sigmoid transfer function, whereas the output neuron used a linear transfer function [32]. During the training process, the first subset is randomly dividing into a training sample, a cross-validation sample, and a test sample (70%, 15%, and 15%, respectively). The connections weights were initialized randomly, and they were tuned using the Levenberg-Marquardt algorithm [80], which is one of the most widely used functions for time series network prediction and training [81]. It is essential to mention that the training was performed using the series-parallel architecture, and the test sample was evaluated with the architecture mentioned above. The test sample is randomly selected, so the sense of temporality is lost. Once the NARX was fitted in a series-parallel configuration, it was converted to parallel architecture, and the test close-loop (second subset) was evaluated accordingly. This model can perform forecasting for several time steps ahead; hence, the predicted outputs

FIGURE 3: Scheme of NARX models.

(at previous steps) constitute a real-time series as well as the test close-loop subset.

2.8. Performance Measures

2.8.1. Nash–Sutcliffe Efficiency Coefficient (NSE). The NSE is widely used to evaluate the performance of hydrological models. NSE is even better than other metrics, such as the coefficient of determination. However, it is susceptible to extreme values since it makes a sum over the square values of the differences between the observed and the predicted values [82]. This index is defined by equation (5):

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}, \quad (8)$$

where O_i and P_i are the observed and predicted values in each period, respectively, and $\bar{O}$ is the average of the observed values.

2.8.2. Kling–Gupta Efficiency (KGE). The KGE is a performance measure based on three equally weighted components: variability, linear correlation, and bias ratio, between predicted and observed data. This index is defined by equation (6):

$$KGE = 1 - \sqrt{(\alpha - 1)^2 + (cc - 1)^2 + (\beta - 1)^2}, \quad (9)$$

where α is the variability (the ratio between the standard deviation of predicted over the observed values), cc is the linear correlation between predicted and observed values, and β is the division between the average of predicted over the average of observed values.

2.8.3. Determination of Bias Percentage (PBIAS). The PBIAS determines whether there is a tendency in the values predicted by the model (i.e., if these are higher or lower than the observed values). A positive PBIAS indicates that the model underestimates the predicted variable, while a negative indicates that the variable is overestimating. The optimal value is a PBIAS equal to zero. This index is defined as

$$PBIAS(\%) = \frac{\sum_{i=1}^N (O_i - P_i) * 100}{\sum_{i=1}^N O_i}. \quad (10)$$

2.8.4. Root Mean Square Error (RMSE). The root mean square error is the difference between forecasted values and the observations. The RMSE is always positive, and values closed to zero indicate a perfect fit; also, RMSE is sensitive to outliers. The RMSE is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}}, \quad (11)$$

where N is the length of the time series.

NSE, KGE, and PBIAS can classify the goodness of fit in four categories [83, 84], as shown in Table 2.

These metrics evaluate the goodness of fit of the models in the training subset and the forecasting performance in the testing subset.

3. Results and Discussion

3.1. Multivariate Linear Model. To obtain a multivariate linear model for predicting rainfall for Labrador and

TABLE 2: Classification of the goodness of fit and performance.

Goodness of fit	NSE	KGE	PBIAS
Very good	>0.75	>0.90	(−10, 10)
Good	(0.65, 0.75]	(0.75, 0.9]	(−15, −10] $\vee$ [10, 15)
Satisfactory	[0.50, 0.65]	[0.50, 0.75]	(−25, −15] $\vee$ [15, 25)
Unsatisfactory	<0.50	<0.50	< −25 $\vee$ > 25

Chirimachay, dataset from January 1964 to December 2015 is used. Each of the fifty models was fitted with $\lambda = 10^d$, with $d = \{-2, -1.9, -1.8, \dots, 9.9, 10\}$. The fitted models obtained for each λ that produced the best performance were selected. NSE metric applied in training and testing sets with the selected models for both ridge regression and lasso is depicted in Figure 4.

For both ridge regression and lasso, a tendency is evident. As lags increase, the NSE grows as well in the training set. However, for Labrado station (Figure 4(a)), the performance grows as lags increase to 18 where the performance in the test set fell. The models fitting the best, for both ridge and lasso, were obtained with around 16 lags. In the models with 16 lags, the performance of both was the same, while the fit was better for the ridge method. Despite having used regularization methods, the overfitting raised from models that used 18 lags or more, where it is clear that the fit increases, but the performance decreases. In the Chirimachay station (Figure 4(b)), the behavior was similar to Labrado. The best models for both ridge and lasso were with around 18 lags. In these models, ridge was better than lasso for both fit and performance. Again, the overfitting problem raised from the model with 19 lags. The top five of the predictors is shown in Appendix 1. These predictors were ranked by the absolute value of the β_j in each fitted model.

In Appendix 1, it can be seen that, for Labrado station, the predictor Niño 1+2 lagged 1 period (Niño 1+2_1) appears as the most influential index in linear models defined by the lasso method. On the contrary, for the ridge method, Niño 1+2 appears in the first six lagged models; hurricane activity lagged 7 periods (HURR_7) and become the most influential from seven to eleven lagged models, and the remaining models are influencing the same rainfall signal with a lag of 12 months. For the Chirimachay station, Niño 1+2 was the most crucial variable because it always appears as the first predictor of the model obtained by lasso. Meanwhile, for ridge regression, the same rainfall (with a lag of one month) is the most influential predictor in models that considered until eleven lags. However, from here (models with lags from twelve to twenty-four months), the same rainfall with lags of 11 and 12 months became the most important predictor. Another essential predictor indexes are as follows: North Pacific Oscillation lagged 1 period (NP_1), and Sahel rainfall lagged 3 or 5 periods (SAHELRAIN_3 and SAHELRAIN_5), which appears in the majority of the models for both lasso and ridge regression model regardless of the station. It is worth noting that when we refer to the predictors, for example, Niño 1+2 lagged 1 period, the true lag is 13 periods since we initially induced a lag of 12 periods in the predictors with respect to the rainfall variable in the initial database. This naturally does not apply to delayed

rainfall variables. ENSO indexes have a strong influence on Ecuadorian rainfall [43]. They found that significant predictors for rainfall come mainly from the tropical Pacific sea surface temperature, especially from ENSO events. This statement matches with our finding on the Niño 1+2 and North Pacific Oscillation indexes. Overall, the models obtained for both Labrado and Chirimachay were “unsatisfactory” according to Table 2, thus implying the need to move in nonlinear models to improve the performance.

3.2. Random Forest Model. The RF model of 12 lags had the best performance in Labrado station. However, the testing displays (Table 3) low values of NSE and KGE; thus, these are classified as “Unsatisfactory”. On the contrary, PBIAS is “very good” value, and the RMSE is relatively high. In the test, the close-loop values of NSE, KGE, and RMSE are bad; only PBIAS is better relative to the test subset. The Chirimachay station presents the better RF model with 12 lags similar to Labrado. Thus, in the testing, NSE and KGE show “unsatisfactory” values and PBIAS is “very good”. In the test close-loop subset, the metrics NSE, KGE, and RMSE show low performance and only PBIAS is “very good”.

The RF models show poor results of forecasting (Figure 5); thus, NSE and KGE have low values, and those are unsatisfactory. The model has difficulty in forecasting high and low extreme values; the difference between observed and forecasting is more evident in January and July for both stations; in general, RF shows a low forecasting performance although RF is better than LM.

Also, Chen et al. [85] found RF performing better than LM for drought forecasting. However, extreme events are not accurate. In concordance with [85], the RF models in the peaks in rainfall are not accurate. Rainfall in Labrado and Chirimachay stations shows two high peaks in the year, in April and October (Figure 2). For this reason, low values of monthly forecasting are observed in the wet and dry seasons.

3.3. Support Vector Machine Model. The SVM model with 3 lags presents a better performance for Labrado station; however, for the test, the statistics are weak, and the NSE and KGE show “unsatisfactory” values; however, PBIAS is “very good,” and the RMSE is relatively high. In the test close-loop, the metrics for evaluating the model worsen; only RMSE shows a little improvement. Therefore, the number of lags does not have a strong influence in the model, SVM models with 3, 6, 9, and 12 lags have very similar values of NSE, KGE, and PBIAS, and the number of lags is not significantly different. In the Chirimachay station, the SVM model with 12 lags exhibits better performance. However, in the test subset, the NSE and KGE

FIGURE 4: NSE metric for models obtained by ridge regression and lasso: (a) Labrado; (b) Chirimachay.

TABLE 3: Results for RF models.

Basin	Lags	Subset	NSE	KGE	PBIAS	RMSE
Labrado	12	Test (Ts)	0.24	0.21	-2.75	39.97
		Test close-loop	0.04	0	1.48	40.41
Chirimachay	12	Test (Ts)	0.25	0.26	-2.37	47.18
		Test close-loop	0.09	0.06	6.1	39.61

FIGURE 5: Continued.

FIGURE 5: Comparison between predicted and observed values: (a) observed vs forecasted subset for Labrado; (b) observed vs forecasted Chirimachay; (c) forecasting for Labrado in 2015; (d) forecasting for Chirimachay in 2015.

present very low “unsatisfactory” values, only PBIAS is “very good,” and RMSE is high. The test close-loop shows that (Table 4) low values of NSE and KGE are “unsatisfactory,” PBIAS is “very good,” and RMSE is relatively high. The SVM model for Labrado and Chirimachay are similar; in other words, in the test and test close-loop shows low values than the test. These values are poor and classified as “unsatisfactory;” also, both models have “very good” PBIAS in test and test close-loop.

Figure 6 shows SVM model forecasted vs. observed data. The model shows difficulty in forecasting rainfall peaks mainly in January and September, where significant discrepancies are observed in both stations. However, the Chirimachay station displays a lower performance. Although SVM is one of most accurate rainfall prediction models [12], it failed to predict extreme rainfall [12]; for this reason, SVM does not show good forecasting in Labrado and Chirimachay, and both shows several low and high peaks in rainfall (Figure 6).

3.4. Recurrent Neural Nets with Exogenous Inputs (NARX).

A different number of hidden neurons were evaluated for its ability to generate accurate model outputs. A hidden layer formed by 50 neurons with sigmoid transfer function and a single output neuron with a linear function provided the most effective network architecture. For the Labrado station, 3 lags in the input were necessary to produce the best model and 6 lags for Chirimachay. Table 5 shows the performance of the best models.

According to Table 5, for Labrado basin, the goodness of fit (evaluated in train, cross-validation, and test samples) is “satisfactory” in accordance with coefficient NSE, “good” in agreement with KGE, and “very good” according to PBIAS. Likewise, the performance (evaluated in the test close-loop) is “satisfactory” for both NSE and KGE and “good” for

PBIAS. For the Chirimachay station, the goodness of fit was also “satisfactory” for both NSE and KGE and “very good” following PBIAS. Finally, the performance was “satisfactory” in agreement with NSE, KGE, and PBIAS. As an overall evaluation, the model fitted for Labrado was “good,” and the model for Chirimachay was “satisfactory.” Figure 7 shows the predicted versus real rainfall in the training subset for Labrado (a) and Chirimachay (b). Besides, the predicted and real rainfall in the test close-loop is presented as time series (c) for Labrado and (d) for Chirimachay.

Figure 7 shows that, for the training subset, the model reached a perfect fit for some data points (points over the line), but for other data points, the predictions were weak (Figures 7(a) and 7(b)). Labrado station shows problems mainly in peaks of rainfall where overestimation and underestimation are present (Figure 7(c)). On the contrary, Chirimachay station shows better forecasting than Labrado, and the tendency is better captured from months 1 to 8, and from 9 to 12, it shows lousy forecasting (Figure 7(d)).

In the test close-loop data points (second dataset), the predictions were “very good” for the first six months, and after that, the performances decrease. It makes sense because the forward predictions are based on predictions from previous steps. Therefore, previous forecast errors affect subsequent predictions. All prediction models depend on long-term data from past time [81]. This is consistent with our study because it works with a long-time series of information and also because the NARX models present good forecasts. However, the prediction quality decrease when the times ahead increase.

The overall results show a satisfactory performance of the NARX model. However, at a lower time scale and with local exogenous variables, the NARX model has shown excellent performance in other studies as well [40, 41]. Also, some

TABLE 4: Results for SVM models.

Basin	Lags	Subset	NSE	KGE	PBIAS	RMSE
Labrado	12	Test (Ts)	0.24	0.23	-0.94	39.85
		Test close-loop	-0.03	-0.03	1.39	41.85
Chirimachay	6	Test (Ts)	0.19	0.31	-0.53	48.93
		Test close-loop	0.09	0.25	0.93	39.65

FIGURE 6: Comparison between predicted and observed values: (a) observed vs forecasted subset for Labrado; (b) observed vs forecasted Chirimachay; (c) forecasting for Labrado in 2015; (d) forecasting for Chirimachay in 2015.

studies, such as [81, 86], show the suitable way of NARX model for getting good predictions on problems involving long-term dependencies. Nevertheless, these studies use few predictors; hence, their performance could improve by including more exogenous variables like our study.

Therefore, NARX models are presented as a good alternative for the prediction of some hydrometeorological variables with various temporal scales because it takes advantages of the information of relevant exogenous variables.

TABLE 5: Results for NARX net models.

Basin	Model	Subset	NSE	KGE	PBIAS	RMSE
Labrado	3 lags, 50 hidden neurons	Train (Tr)	0.57	0.77	-0.50	31.20
		Cross-validation (Cv)	0.38	0.70	0.60	30.83
		Test (Ts)	0.54	0.77	-3.00	30.02
		Test close-loop	0.57	0.51	12.30	47.74
Chirimachay	6 lags, 50 hidden neurons	Train (Tr)	0.52	0.75	2.60	42.72
		Cross-validation (Cv)	0.53	0.72	-5.20	46.20
		Test (Ts)	0.54	0.69	5.10	36.69
		Test close-loop	0.61	0.61	22.20	36.76

FIGURE 7: Comparison between predicted and observed values: (a) training subset for Labrado; (b) training subset for Chirimachay; (c) forecasting for Labrado in 2015; (d) forecasting for Chirimachay in 2015.

APPENDIX

Lags	Model	Labrado					Chirimachay				
		1	2	3	4	5	1	2	3	4	5
L0	Ridge	Niño 1+2	Niño 3	ONI	WP	HURR	Niño 1+2	Niño 3	SOLAR	TNI	TSA
	Lasso	Niño 3	Niño 1+2	ONI	WP	NTA	Niño 3	Niño 1+2	ONI	SOLAR	TNI
L1	Ridge	Niño 1+2_1	NP_1	ONI	HURR_1	Rainfall_1	Rainfall_1	Niño 1+2_1	NP_1	HURR_1	TSA_1
	Lasso	Niño 1+2_1	NP_1	HURR_1	Rainfall_1	TSA_1	Niño 1+2_1	Rainfall_1	NP_1	HURR_1	TSA_1
L2	Ridge	Niño 1+2_1	NP_1	Rainfall_1	Niño 1+2_2	HURR_1	Rainfall_1	NP_1	Niño 1+2_1	HURR_1	NAO_2
	Lasso	Niño 1+2_1	NP_1	HURR_1	Rainfall_1	GMSST_2	Niño 1+2_1	Rainfall_1	NP_1	HURR_1	NAO_2
L3	Ridge	Niño 1+2_1	NP_1	Rainfall_1	Rainfall_3	SAHELRAIN_3	Rainfall_1	NP_1	Rainfall_3	Niño 1+2_1	NAO_2
	Lasso	Niño 1+2_1	NP_1	Rainfall_1	HURR_1	Rainfall_3	Niño 1+2_1	Rainfall_1	NP_1	NAO_2	Rainfall_3
L4	Ridge	NP_1	Niño 1+2_1	Rainfall_3	Rainfall_1	SAHELRAIN_3	Rainfall_1	Rainfall_3	NP_1	Niño 1+2_1	NAO_2
	Lasso	Niño 1+2_1	Rainfall_1	Niño 1+2_4	NP_1	Rainfall_3	Niño 1+2_1	Rainfall_1	NP_1	HURR_1	Rainfall_3
L5	Ridge	Niño 1+2_1	Rainfall_1	Rainfall_3	NP_1	SAHELRAIN_3	Rainfall_1	Rainfall_3	NP_1	Niño 1+2_1	NAO_2
	Lasso	Niño 1+2_1	Niño 1+2_5	Rainfall_3	Rainfall_1	Niño 4_5	Niño 1+2_1	Rainfall_1	NP_1	Rainfall_3	NAO_2
L6	Ridge	Niño 1+2_1	Rainfall_3	Rainfall_3	SAHELRAIN_3	NP_1	Rainfall_1	Rainfall_6	Niño 1+2_1	Rainfall_3	NP_1
	Lasso	Niño 1+2_1	Niño 1+2_5	Rainfall_3	Rainfall_1	Niño 4_5	Niño 1+2_1	Rainfall_1	NP_1	Rainfall_6	Rainfall_3
L7	Ridge	HURR_7	Rainfall_1	Niño 1+2_1	Rainfall_3	NP_1	Rainfall_1	HURR_7	Rainfall_3	Rainfall_6	Niño 1+2_1
	Lasso	Niño 1+2_1	Niño 1+2_5	HURR_7	Rainfall_1	Rainfall_3	Niño 1+2_1	Rainfall_1	NP_1	HURR_7	Rainfall_3
L8	Ridge	HURR_7	Niño 1+2_1	Rainfall_1	NP_1	Rainfall_3	Rainfall_1	HURR_7	Rainfall_3	Rainfall_6	Rainfall_8
	Lasso	Niño 1+2_1	Niño 1+2_5	HURR_7	Rainfall_3	Rainfall_1	Niño 1+2_1	Rainfall_1	NP_1	HURR_7	Rainfall_8
L9	Ridge	HURR_7	SAHELRAIN_3	Rainfall_3	Rainfall_1	NP_1	Rainfall_1	HURR_7	Rainfall_6	Rainfall_8	Rainfall_3
	Lasso	Niño 1+2_1	Niño 1+2_5	HURR_7	Rainfall_3	Rainfall_1	Niño 1+2_1	Rainfall_1	HURR_7	NP_1	Rainfall_8
L10	Ridge	HURR_7	Rainfall_1	Niño 1+2_1	SAHELRAIN_3	NP_1	Rainfall_1	HURR_7	Rainfall_6	Niño 1+2_1	Rainfall_3
	Lasso	Niño 1+2_1	HURR_7	Rainfall_1	Rainfall_3	BEST_9	Niño 1+2_1	Rainfall_1	NP_1	HURR_7	Rainfall_8
L11	Ridge	HURR_7	Rainfall_3	SAHELRAIN_3	Rainfall_1	NP_1	Rainfall_1	Rainfall_11	HURR_7	Rainfall_6	Rainfall_8
	Lasso	Niño 1+2_1	HURR_7	Rainfall_3	Rainfall_1	SOI_11	Niño 1+2_1	Rainfall_1	Rainfall_11	HURR_7	NP_1
L12	Ridge	Rainfall_12	HURR_7	Rainfall_3	SAHELRAIN_3	NP_1	Rainfall_12	Rainfall_1	Rainfall_11	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	SOI_11	Rainfall_3	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1
L13	Ridge	Rainfall_12	HURR_7	Rainfall_3	NP_1	SAHELRAIN_3	Rainfall_12	Rainfall_1	Rainfall_11	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	Rainfall_3	SOI_11	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1
L14	Ridge	Rainfall_12	HURR_7	Rainfall_3	NP_1	Rainfall_1	Rainfall_1	Rainfall_12	Rainfall_11	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	Rainfall_3	SOI_11	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1
L15	Ridge	Rainfall_12	HURR_7	Rainfall_3	NP_1	Niño 1+2_1	Rainfall_12	Rainfall_11	Rainfall_1	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	NP_1	Rainfall_1	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	HURR_7
L16	Ridge	Rainfall_12	Rainfall_3	HURR_7	SAHELRAIN_3	SAHELRAIN_5	Rainfall_11	Rainfall_12	Rainfall_1	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	Rainfall_3	Rainfall_1	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1
L17	Ridge	Rainfall_12	HURR_7	NP_1	Niño 1+2_1	Rainfall_3	Rainfall_12	Rainfall_1	Rainfall_11	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Rainfall_12	HURR_7	NP_1	Rainfall_1	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1
L18	Ridge	Rainfall_12	HURR_7	Rainfall_3	NP_1	Niño 1+2_1	Rainfall_12	Rainfall_11	Rainfall_1	HURR_7	Rainfall_6
	Lasso	Niño 1+2_1	Niño 1+2_18	Rainfall_12	NP_1	HURR_7	Niño 1+2_1	Rainfall_1	Rainfall_12	Rainfall_11	NP_1

TABLE : Continued.

Lags	Model	Labrado					Chirimachay				
		1	2	3	4	5	1	2	3	4	5
L19	Ridge	Rainfall_3	Rainfall_12	SAHELRAIN_3	HURR_7	SAHELRAIN_5	Rainfall_12	Rainfall_11	Rainfall_1	HURR_19	Rainfall_6
	Lasso	Niño 1 + 2_1	Niño 1 + 2_18	Rainfall_12	NP_1	Rainfall_1	Niño 1 + 2_1	Rainfall_1	HURR_19	NP_1	Rainfall_12
L20	Ridge	Rainfall_12	Rainfall_3	HURR_7	HURR_19	SAHELRAIN_3	Rainfall_12	Rainfall_11	Rainfall_1	HURR_19	Rainfall_17
	Lasso	Niño 1 + 2_18	Niño 1 + 2_1	Rainfall_12	NP_1	Rainfall_1	Niño 1 + 2_1	Rainfall_1	HURR_19	Rainfall_12	Rainfall_11
L21	Ridge	Rainfall_12	Rainfall_3	HURR_7	SAHELRAIN_3	HURR_19	Rainfall_11	Rainfall_12	Rainfall_1	HURR_19	Rainfall_6
	Lasso	Niño 1 + 2_1	Rainfall_3	Rainfall_12	NP_1	HURR_7	Niño 1 + 2_1	Rainfall_1	HURR_19	Rainfall_11	Rainfall_12
L22	Ridge	Rainfall_12	Rainfall_3	HURR_7	HURR_19	SAHELRAIN_3	Rainfall_11	Rainfall_12	Rainfall_1	HURR_19	Rainfall_6
	Lasso	Niño 1 + 2_1	Rainfall_12	Niño 1 + 2_18	NP_1	Rainfall_3	Niño 1 + 2_1	Rainfall_1	HURR_19	Niño 3_16	NP_1
L23	Ridge	Rainfall_3	Rainfall_12	SAHELRAIN_3	HURR_7	AO_14	Rainfall_11	Rainfall_12	Rainfall_1	HURR_19	Rainfall_6
	Lasso	Niño 1 + 2_1	Rainfall_3	Rainfall_12	Niño 1 + 2_18	NP_1	Niño 1 + 2_1	Rainfall_1	Rainfall_11	Rainfall_12	HURR_19
L24	Ridge	Rainfall_3	SAHELRAIN_3	Rainfall_12	AO_14	SAHELRAIN_5	Rainfall_11	Rainfall_12	Rainfall_1	HURR_19	Rainfall_6
	Lasso	Niño 1 + 2_1	Rainfall_3	Rainfall_12	Niño 1 + 2_18	NP_1	Niño 1 + 2_1	Rainfall_1	HURR_19	Rainfall_11	Rainfall_12

4. Conclusions

In this study, linear and nonlinear approaches are evaluated for rainfall forecasting. For the linear models, the climatic indices: Niño 1+2, Sahel rainfall, hurricane activity, North Pacific Oscillation, and the same delayed rainfall signal, were the most influential in the prediction of rainfall. Even for models with predictors with delays higher than or equal to 12 months, the same rainfall signal that delayed 12 months was the most significant. Thus, it shows the seasonality present in the rainfall signal, which expected similar measures for the same month every year. However, the linear models did not capture well the rainfall dynamics, and hence, the evaluations were “unsatisfactory.” Linear models reached the best performances with predictors lagging around 16 or 18 periods plus an initial delay of 12 months between the predictors and the rainfall variable. The RF and SVM showed better forecasting than LM in both stations, but this model presents inaccuracy in low and high peaks of rainfall. The SVM exhibits better performance than RF; nevertheless, in general, both display poor forecasting performance in Labrado and Chirimachay.

On the contrary, the nonlinear autoregressive network with exogenous inputs was also evaluated. Considering that the behavior of the rainfall is chaotic and highly nonlinear, the NARX networks had a better performance, obtaining models considered “good” for Labrado station and “satisfactory” for Chirimachay station. The best NARX networks were reported, which had 50 neurons in the hidden layer, and the delays were multiples of 3 periods (3 for Labrado and 6 for Chirimachay). In the prediction of the 12 months ahead, the models almost always follow the real changes in the rainfall signal. For the first 6 months, the models are very accurate in the prediction. Nevertheless, as we continue the prediction ahead, the performance decreases.

The forecasting shows several problems for predicting peaks; this was more evident in the Labrado station. These results could be attributed to the complex behavior of weather in Andean regions, which has several external and local influences affecting the dynamic of rainfall processes. Thus, the future challenges on rainfall prediction must focus on capturing these peaks and identify the triggers of the precipitation events. These improvements in rainfall predictions are essential to generate plans for suitable decision-making in water resources management.

Appendix

Predictors order

Data Availability

The times series data of precipitation and climatic indices of both stations used to support the findings of this study are included within the supplementary information files.

Disclosure

The participation of A.V.-P. is made in the context of his doctoral program in water resources, offered by Universidad

de Cuenca, Escuela Politécnica Nacional, and Universidad Técnica Particular de Loja.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors thank the INAMHI for the information provided. This work was supported by the University of Cuenca through its Research Department (DIUC) via the projects “Evaluación del Riesgo de sequías en cuencas andinas reguladas influenciadas por la variabilidad climática y cambio climático” and “Estudio de las condiciones climáticas de América del Sur que producen las sequías en el Ecuador continental”.

Supplementary Materials

The files DB_Chirimachay_station.txt and DB_Labrado_station.txt are organized in the following way: firstly, the dates, secondly, the time series of the 27 climatic indices, and thirdly, the time series of the rain station. (*Supplementary Materials*)

References

- [1] E. A. B. Eltahir and R. L. Bras, “Precipitation recycling,” *Reviews of Geophysics*, vol. 34, no. 3, pp. 367–378, 1996.
- [2] Q. Sun, C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K. L. Hsu, “A review of global precipitation data sets: data sources, estimation, and intercomparisons,” *Reviews of Geophysics*, vol. 56, no. 1, pp. 79–107, 2018.
- [3] C. Kidd and G. Huffman, “Global precipitation measurement,” *Meteorological Applications*, vol. 18, no. 3, pp. 334–353, 2011.
- [4] M. Xinggang, J. Wenxiong, Z. Guofeng et al., “Stable isotope composition of precipitation at different elevations in the monsoon marginal zone,” *Quaternary International*, vol. 493, pp. 86–95, 2018.
- [5] M. N. French, W. F. Krajewski, and Y. R. R. Cuykendall, “Rainfall forecasting in space and time using a neural network,” *Journal of Hydrology*, vol. 137, no. 1–4, pp. 1–31, 1992.
- [6] C. Daly, D. R. Conklin, and M. H. Unsworth, “Local atmospheric decoupling in complex topography alters climate change impacts,” *International Journal of Climatology*, vol. 30, no. 12, pp. 1857–1864, 2010.
- [7] R. Céleri and J. Feyen, “The hydrology of tropical andean ecosystems: importance, knowledge status, and perspectives,” *Mountain Research and Development*, vol. 29, no. 4, pp. 350–355, 2009.
- [8] W. Buytaert, V. Iniguez, and B. D. Bièvre, “The effects of afforestation and cultivation on water yield in the Andean páramo,” *Forest Ecology and Management*, vol. 251, no. 1–2, pp. 22–30, 2007.
- [9] W. Buytaert, R. Céleri, B. De Bièvre et al., “Human impact on the hydrology of the Andean páramos,” *Earth-Science Reviews*, vol. 79, no. 1–2, pp. 53–72, 2006.
- [10] A. DDubey, “Artificial neural network models for rainfall prediction in pondicherry,” *International Journal of Computer Applications*, vol. 120, no. 3, pp. 30–35, 2015.

- [11] N. Mishra, H. K. Soni, H. K. Soni, S. Sharma, and A. K. Upadhyay, "Development and analysis of artificial neural network models for rainfall prediction by using time-series data," *International Journal of Intelligent Systems and Applications*, vol. 10, no. 1, pp. 16–23, 2018.
- [12] A. M. Bagirov and A. Mahmood, "A comparative assessment of models to predict monthly rainfall in Australia," *Water Resources Management*, vol. 32, no. 5, pp. 1777–1794, 2018.
- [13] J.-H. Seo, Y. H. Lee, and Y.-H. Kim, "Feature selection for very short-term heavy rainfall prediction using evolutionary computation," *Advances in Meteorology*, vol. 2014, pp. 1–15, 2014.
- [14] A. Givati, M. Housh, Y. Levi, D. Paz, I. Carmona, and E. Becker, "The advantage of using international multimodel ensemble for seasonal precipitation forecast over Israel," *Advances in Meteorology*, vol. 2017, pp. 1–11, 2017.
- [15] A. Schepen, T. Zhao, Q. J. Wang, and D. E. Robertson, "A Bayesian modelling method for post-processing daily sub-seasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments," *Hydrology and Earth System Sciences*, vol. 22, no. 2, pp. 1615–1628, 2018.
- [16] A. Schepen, Q. J. Wang, and D. E. Robertson, "Seasonal forecasts of Australian rainfall through calibration and bridging of coupled GCM outputs," *Monthly Weather Review*, vol. 142, no. 5, pp. 1758–1770, 2014.
- [17] S. Khajehei and H. Moradkhani, "Towards an improved ensemble precipitation forecast: a probabilistic post-processing approach," *Journal of Hydrology*, vol. 546, pp. 476–489, 2017.
- [18] A. Möller, A. Lenkoski, and T. L. Thorarinsdottir, "Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 673, pp. 982–991, 2013.
- [19] A. Schepen and Q. J. Wang, "Ensemble forecasts of monthly catchment rainfall out to long lead times by post-processing coupled general circulation model output," *Journal of Hydrology*, vol. 519, pp. 2920–2931, 2014.
- [20] J. Byakatonda, B. P. Parida, P. K. Kenabatho, and D. B. Moalafhi, "Prediction of onset and cessation of austral summer rainfall and dry spell frequency analysis in semiarid Botswana," *Theoretical and Applied Climatology*, vol. 135, no. 1–2, pp. 101–117, 2019.
- [21] M. Dounia, D. Sabri, and D. Yassine, "Rainfall-rain off modeling using artificial neural network," *APCBEE Procedia*, vol. 10, pp. 251–256, 2014.
- [22] H. Mislán, S. Hardwinarto, Y. Sumaryono, and M. Aipassa, "Rainfall monthly prediction based on artificial neural network: a case study in tenggarong station, East Kalimantan-Indonesia," *Procedia Computer Science*, vol. 59, pp. 142–151, 2015.
- [23] V. K. Dabhi and S. Chaudhary, "Hybrid wavelet-postfix-GP model for rainfall prediction of anand region of India," *Advances in Artificial Intelligence*, vol. 2014, pp. 1–11, 2014.
- [24] Z. Chao, F. Pu, Y. Yin, B. Han, and X. Chen, "Research on real-time local rainfall prediction based on MEMS sensors," *Journal of Sensors*, vol. 2018, pp. 1–9, 2018.
- [25] C. Cai, J. Wang, and Z. Li, "Improving TIGGE precipitation forecasts using an SVR ensemble approach in the huaihe river basin," *Advances in Meteorology*, vol. 2018, pp. 1–15, 2018.
- [26] S.-H. Moon, Y.-H. Kim, Y. H. Lee, and B.-R. Moon, "Application of machine learning to an early warning system for very short-term heavy rainfall," *Journal of Hydrology*, vol. 568, pp. 1042–1054, 2019.
- [27] M. Min, C. Bai, J. Guo et al., "Estimating summertime precipitation from himawari-8 and global forecast system based on machine learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2557–2570, 2019.
- [28] M. Ali, R. Prasad, Y. Xiang, and Z. M. Yaseen, "Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts," *Journal of Hydrology*, vol. 58, Article ID 124647, 2020.
- [29] C. Okonkwo, "An advanced review of the relationships between Sahel precipitation and climate indices: a wavelet approach," *International Journal of Atmospheric Sciences*, vol. 2014, pp. 1–11, 2014.
- [30] A. Ali, "Nonlinear multivariate rainfall-stage model for large wetland systems," *Journal of Hydrology*, vol. 374, no. 3–4, pp. 338–350, 2009.
- [31] F.-J. Chang, L.-C. Chang, C.-W. Huang, and I.-F. Kao, "Prediction of monthly regional groundwater levels through hybrid soft-computing techniques," *Journal of Hydrology*, vol. 541, pp. 965–976, 2016.
- [32] S. M. Guzman, J. O. Paz, and M. L. M. Tagert, "The use of NARX neural networks to forecast daily groundwater levels," *Water Resources Management*, vol. 31, no. 5, pp. 1591–1603, 2017.
- [33] A. Wunsch, T. Liesch, and S. Broda, "Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX)," *Journal of Hydrology*, vol. 567, pp. 743–758, 2018.
- [34] F. Previdi and M. Lovera, "Identification of parametrically-varying models for the rainfall-runoff relationship in urban drainage networks," *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 1768–1773, 2009.
- [35] B. A. Amisigo, N. van de Giesen, C. Rogers, W. E. I. Andah, and J. Friesen, "Monthly streamflow prediction in the volta basin of west africa: a SISO narmax polynomial modelling," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 33, no. 1–2, pp. 141–150, 2008.
- [36] M. Ebrahim, A. Ahmadian, and Y. F. Sadat, "GeoResJ hybrid darima-narx model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic oscillation (NAO) and rainfall data," *GeoResJ*, vol. 13, pp. 9–16, 2017.
- [37] F.-J. Chang, P.-A. Chen, Y.-R. Lu, E. Huang, and K.-Y. Chang, "Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control," *Journal of Hydrology*, vol. 517, pp. 836–846, 2014.
- [38] H. Kim, H. Keum, and K. Han, "Real-time urban inundation prediction combining hydraulic and probabilistic methods," *Water*, vol. 11, no. 2, pp. 293–318, 2019.
- [39] T. Nanda, B. Sahoo, H. Beria, and C. Chatterjee, "A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products," *Journal of Hydrology*, vol. 539, pp. 57–73, 2016.
- [40] P. Benevides, J. Catalao, and G. Nico, "Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors," *Remote Sensing*, vol. 11, no. 8, p. 966, 2019.
- [41] Z. Rahimi, H. Z. Mohd Shafri, and M. Norman, "A GNSS-based weather forecasting approach using nonlinear autoregressive approach with exogenous input (NARX)," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 178, pp. 74–84, 2018.
- [42] E. Pisoni, M. Farina, C. Carnevale, and Y. L. Piroddi, "Forecasting peak air pollution levels using NARX models,"

- Engineering Applications of Artificial Intelligence*, vol. 22, no. 4-5, pp. 593–602, 2009.
- [43] L. B. de Guenni, M. García, Á. G. Muñoz et al., “Predicting monthly precipitation along coastal Ecuador: ENSO and transfer function models,” *Theoretical and Applied Climatology*, vol. 129, no. 3-4, pp. 1059–1073, 2017.
- [44] D. E. Mendoza, E. P. Samaniego, E. A. Pacheco, V. M. Carrillo, and y D. R. Ochoa-Tocachi, “Stochastic rainfall forecasting for high tropical andean regions,” 2017.
- [45] R. Céleri, L. Campozaño, and y A. Avilés, “Integrated waterresources management in an Andean mountain basin: thecase of the Machángara river,” in *Nexus Outlook: Assessing Resource Use Challenges in the Water, Energy and Food Nexus*, M. Al-aidi and L. Ribbe, Eds., pp. 80–97, Research Gate, Berlin, Germany, 2016.
- [46] B. C. S. Hansen, D. T. Rodbell, G. O. Seltzer, B. León, K. R. Young, and M. Abbott, “Late-glacial and Holocene vegetational history from two sites in the western Cordillera of Southwestern Ecuador,” *Palaeogeography, Palaeoclimatology, Palaeoecology*, vol. 194, no. 1–3, pp. 79–108, 2003.
- [47] V. Vanacker, G. Govers, S. Barros, J. Poesen, and J. Deckers, “The effect of short-term socio-economic and demographic change on landuse dynamics and its corresponding geomorphic response with relation to water erosion in a tropical mountainous catchment, Ecuador,” *Landscape Ecology*, vol. 18, no. 1, pp. 1–15, 2003.
- [48] R. Celleri, P. Willems, W. Buytaert, and J. Feyen, “Space-time rainfall variability in the Paute basin, Ecuadorian Andes,” *Hydrological Processes*, vol. 21, no. 24, pp. 3316–3327, 2007.
- [49] J. C. H. Chiang and D. J. Vimont, “Analogous pacific and atlantic meridional modes of tropical atmosphere-ocean variability,” *Journal of Climate*, vol. 17, no. 21, pp. 4143–4158, 2004.
- [50] D. B. Enfield, A. M. Mestas-Núñez, and P. J. Trimble, “The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S,” *Geophysical Research Letters*, vol. 28, no. 10, pp. 2077–2080, 2001.
- [51] S. Zhou, A. J. Miller, J. Wang, and J. K. Angell, “Trends of NAO and AO and their associations with stratospheric processes,” *Geophysical Research Letters*, vol. 28, no. 21, pp. 4107–4110, 2001.
- [52] C. A. Smith and P. D. Sardeshmukh, “The effect of ENSO on the intraseasonal variance of surface temperatures in winter,” *International Journal of Climatology*, vol. 20, no. 13, pp. 1543–1557, 2000.
- [53] C. Penland and L. Matrosova, “Prediction of tropical atlantic sea surface temperatures using linear inverse modeling,” *Journal of Climate*, vol. 11, no. 3, pp. 483–496, 1998.
- [54] G. S. Lehmiller, T. B. Kimberlain, and J. B. Elsner, “Seasonal prediction models for North Atlantic basin Hurricane location,” *Monthly Weather Review*, vol. 125, no. 8, pp. 1780–1791, 1997.
- [55] K. Wolter and M. S. Timlin, “El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext),” *International Journal of Climatology*, vol. 31, no. 7, pp. 1074–1087, 2011.
- [56] A. G. Barnston and R. E. Livezey, “Classification, seasonality and persistence of low-frequency atmospheric circulation patterns,” *Monthly Weather Review*, vol. 115, no. 6, pp. 1083–1126, 1987.
- [57] K. Trenberth, *The Climate Data Guide: Nino SST Indices (Nino 1 + 2, 3, 3.4, 4; ONI and TNI)*, National Center for Atmospheric Research, Boulder, CO, USA, 2016.
- [58] J. C. Rogers, “The North Pacific oscillation,” *Journal of Climatology*, vol. 1, no. 1, pp. 39–57, 1981.
- [59] J.-Y. Yu, H.-Y. Kao, T. Lee, and S. T. Kim, “Subsurface Ocean Temperature indices for central-pacific and eastern-pacific types of el Niño and La niña events,” *Theoretical and Applied Climatology*, vol. 103, no. 3-4, pp. 337–344, 2011.
- [60] N. J. Mantua and S. R. Hare, “The Pacific decadal oscillation,” *Journal of Oceanography*, vol. 58, no. 1, pp. 35–44, 2002.
- [61] R. A. Ebdon, “Notes on the wind flow at 50 mb in tropical and sub-tropical regions in January 1957 and January 1958,” *Quarterly Journal of the Royal Meteorological Society*, vol. 86, no. 370, pp. 540–542, 1960.
- [62] R. J. Reed, W. J. Campbell, L. A. Rasmussen, and D. G. Rogers, “Evidence of a downward-propagating, annual wind reversal in the equatorial stratosphere,” *Journal of Geophysical Research*, vol. 66, no. 3, pp. 813–818, 1961.
- [63] E. S. Nicholson, *A Review of Recent Studies on the Rainfall Regime and Its Interannual Variability*, Vol. 32306, Earth, Ocean, and Atmospheric Sciences Department, Florida State University, Tallahassee, FL, USA, 2013.
- [64] A. J. Troup, “The southern oscillation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 91, no. 390, pp. 490–506, 1965.
- [65] I. Poole, “Understanding solar indices,” *Qst*, vol. 200, pp. 38–40, 2002.
- [66] D. B. Enfield, A. M. Mestas-Núñez, D. A. Mayer, and L. Cid-Serrano, “How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures?” *Journal of Geophysical Research: Oceans*, vol. 104, no. C4, pp. 7841–7848, 1999.
- [67] K. E. Trenberth and D. P. Stepaniak, “Indices of el Niño evolution,” *Journal of Climate*, vol. 14, no. 8, pp. 1697–1701, 2001.
- [68] B. Rajagopalan, Y. Kushnir, and Y. M. Tourre, “Observed decadal midlatitude and tropical Atlantic climate variability,” *Geophysical Research Letters*, vol. 25, no. 21, pp. 3967–3970, 1998.
- [69] C. Wang and D. B. Enfield, “The tropical western Hemisphere warm pool,” *Geophysical Research Letters*, vol. 28, no. 8, pp. 1635–1638, 2001.
- [70] J. M. Wallace and D. S. Gutzler, “Teleconnections in the geopotential height field during the northern Hemisphere winter,” *Monthly Weather Review*, vol. 109, no. 4, pp. 784–812, 1981.
- [71] J. Hansen, R. Ruedy, J. Glascoe, and M. Sato, “GISS analysis of surface temperature change,” *Journal of Geophysical Research: Atmospheres*, vol. 104, no. D24, pp. 30997–31022, 1999.
- [72] J. Hansen, M. Sato, and K. Lo, “Global surface temperature change,” *Global Pipeline Monthly*, vol. 5, no. 2, pp. 1–29, 2009.
- [73] G. James, D. Witten, T. Hastie, and Y. R. Tibshirani, *An Introduction to Statistical Learning*, Vol. 7, Springer, Berlin, Germany, 2000.
- [74] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014.
- [75] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.
- [76] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [77] J. R. Quinlan, “Induction of decision trees,” *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.
- [78] M. Kubat, “A simple machine-learning task,” in *An Introduction to Machine Learning*, M. Kubat, pp. 1–18, Springer International Publishing, Cham, Switzerland, 2017.

- [79] V. N. Vapnik, "Direct methods in statistical learning theory," in *The Nature of Statistical Learning Theory*, V. N. Vapnik, Ed., pp. 225–265, Springer, New York, NY, USA, 2000.
- [80] H. B. Demuth, M. H. Beale, O. De Jess, and Y. M. T. Hagan, *Neural Network Design*, Martin Hagan, Stillwater, OK, USA, 2014.
- [81] T. Lin, B. G. Horne, P. Tino, and C. L. Giles, "Learning long-term dependencies in NARX recurrent neural networks," *IEEE Transactions Neural Networks*, vol. 7, pp. 1329–1338, 1996, <https://ieeexplore.ieee.org/abstract/document/548162>.
- [82] D. R. Legates and G. J. McCabe, "Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation," *Water Resources Research*, vol. 35, no. 1, pp. 233–241, 1999.
- [83] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith, "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations," *Transactions of the American Society of Agricultural and Biological Engineers*, vol. 50, no. 3, pp. 885–900, 2007.
- [84] K. Abbaspour, S. Raghavan, S. A. Vaghefi, M. Faramarzi, and Y. L. Chen, *Integrated Soil and Water Management: Selected Papers from 2016 International SWAT Conference*, MDPI, Basel, Switzerland, 2017.
- [85] J. Chen, M. Li, and y. W. Wang, "Statistical uncertainty estimation using random forests and its application to drought forecast," *Mathematical Problems in Engineering*, vol. 2012, Article ID 915053, 2012.
- [86] H. T. Siegelmann, B. G. Horne, and y. C. L. Giles, "Computational capabilities of recurrent NARX neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 27, no. 2, pp. 208–215, 1997.