Use of nonlinear principal components of CHIRPS precipitation data and ocean-atmospheric variables for streamflow forecasting in an area of scarce data. Case study, Tocaría river basin - Orinoquia Colombiana

5 Sarria-Ospina, Jhon Derly[1]; Ocampo-Marulanda, Camilo[2,3]; Ceron-Aramburo, Lina Maria[1]; Canchala, Teresita[4]; Ferreira, Tiago Alessandro[2]

[1]Research Group TERRANARE, Faculty of Natural Sciences and Engineering, Fundación Universitaria de San Gil, Yopal, 850001, Colombia

[2]Programa de Pós-graduação em Biometria e Estatística Aplicada, Departamento de Estatística e Informática, Universidade
10 Federal Rural de Pernambuco, Recife, 52171-900, Brasil

[3]Water Resources, Engineering and Soil Research Group (IREHISA), School of Natural Resources and Environmental Engineering,   Universidad   del   Valle,   760032,   Cali,   Colombia

[4]Environmental Engineering Program, Faculty of Engineering, Universidad Mariana, Pasto, 520002, Colombia

Correspondence to: Sarria-Ospina, Jhon (jsarria.sgryopal@unisangil.edu.co)

15 **Abstract.** Accurate streamflow forecasting is critical for mitigating the impacts of hydrological extremes and guiding sustainable water resource management, particularly in poorly gauged tropical catchments. This study presents a hybrid forecasting framework that integrates Neural Network Seasonal Autoregressive Integrated Moving Average using exogenous variables (NN-SARIMAX) models with nonlinear principal components (NLPCs) derived from CHIRPS precipitation data, and large-scale ocean–atmosphere indices (macroclimatic variables, MVs). Four monthly models were developed and tested
20 for the Tocaría River basin in the Colombian Orinoquía region: (1) a baseline SARIMA $(4,0,4)(0,0,3)_{12}$ model; (2) SARIMAX with exogenous MVs; (3) NN-SARIMAX with NLPCs; and (4) a hybrid NN-SARIMAX combining both MVs and NLPCs. The hybrid model achieved the best performance with an R^2 of 0.78 during the validation period. These results underscore the effectiveness of integrating local precipitation variability and large-scale climatic drivers to enhance forecast accuracy under data-scarce conditions. The proposed methodology offers a transferable approach for operational forecasting in ungauged or
25 sparsely monitored basins, contributing to early warning systems, drought preparedness, and adaptive water governance in vulnerable tropical regions.

Keywords: monthly forecasting, macroclimatic variables, NLPC, SARIMA, exogenous variables

1 INTRODUCTION

Sustainable management of water resources in tropical basins critically depends on the ability to accurately anticipate
30 streamflow dynamics amid complex hydroclimatic variability. In regions like the Colombian Orinoquía, the interplay of large-scale ocean–atmosphere phenomena; especially the El Niño–Southern Oscillation (ENSO); introduces pronounced

nonstationarity and nonlinearity, challenging traditional time series forecasting methods (Pasquini et al., 2007; Chiew and McMahon, 2002; Yap and Musa, 2023). Streamflow integrates precipitation, evapotranspiration, infiltration, and climatic anomalies across multiple spatial and temporal scales, yet sparse hydrometeorological monitoring limits the characterization

35 of these processes (Tootle et al., 2008; Yao et al., 2020).

Conventional statistical models such as Autoregressive Integrated Moving Average (ARIMA) and its seasonal extension (SARIMA) have been widely applied for streamflow prediction due to their capacity to model temporal dependencies and seasonality under stationarity assumptions (Valipour, 2015; Sirisha et al., 2022). However, their performance diminishes in capturing abrupt, nonlinear fluctuations characteristic of tropical catchments influenced by exogenous climate drivers (Moeeni

40 and Bonakdari, 2017). Consequently, hybrid approaches that couple SARIMA with Artificial Neural Networks (ANNs) have emerged, leveraging ANNs' nonlinear learning capacity and robustness to data scarcity to improve forecasting skill (Rafael et al., 2022; Niu and Feng, 2021; Costa et al., 2023).

In Colombia, ANN-based models incorporating macroclimatic indices; such as sea surface temperature anomalies and ENSO-related indices; have demonstrated substantial improvements in streamflow forecasting accuracy for Andean and Pacific basins

45 (Mesa et al., 2001; Poveda et al., 2002; Velásquez et al., 2010; Cárdenas et al., 2022). Nevertheless, similar studies remain scarce for the Orinoquía region, despite its hydrological sensitivity to both Pacific and Atlantic climatic influences (Builes-Jaramillo et al., 2022). The Tocaría River basin (Orinoquía region), exemplifies this knowledge gap; it is characterized by limited hydrometeorological instrumentation and experiences seasonal extremes of drought and flooding that impact local agriculture and ecosystems (Corporinoquia and Corpoboyacá, 2015a).

50 Building on previous work that successfully employed NLPCA combined with satellite-derived CHIRPS precipitation data to impute missing monthly streamflow records in the Tocaría basin; demonstrating high accuracy across stations with substantial data gaps (Ocampo-Marulanda et al., 2025); this study advances the development of a parsimonious hybrid forecasting framework. The proposed model integrates a SARIMA component with feedforward ANNs and incorporates NLPC derived from CHIRPS data alongside macroclimatic variables representing large-scale atmospheric–oceanic drivers named NN-

55 SARIMAX. This nonlinear dimensionality reduction enhances model generalization in low-data contexts, effectively capturing the concurrent influences of Pacific and Atlantic climate variability.

The research aims to provide an accurate monthly streamflow forecasting tool tailored to the Tocaría River's hydrological regime, supporting adaptive water resource management in a data-limited and climate-vulnerable tropical basin. By addressing key gaps in scientific knowledge and practical decision-making tools, the framework offers scalable potential for application

60 in other similarly under-monitored tropical catchments subject to complex climatic forcing.

2 METHODOLOGY

This section outlines the methodological framework developed for monthly streamflow forecasting in poorly instrumented basins, using the Tocaría River (Colombian Orinoquia) as a case study. The approach integrates seasonal autoregressive modeling with exogenous predictors, nonlinear dimensionality reduction, and multivariate correlation analysis.

As illustrated in Fig 1, the workflow combines three primary data sources—observed streamflow, satellite-derived precipitation (CHIRPS), and ocean–atmosphere macroclimatic variables (MVs)—through preprocessing, imputation, NLPCA-based reduction, and statistical evaluation. Four forecasting models were independently constructed and trained, and subsequently compared using RMSE, R², Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and residual diagnostics.
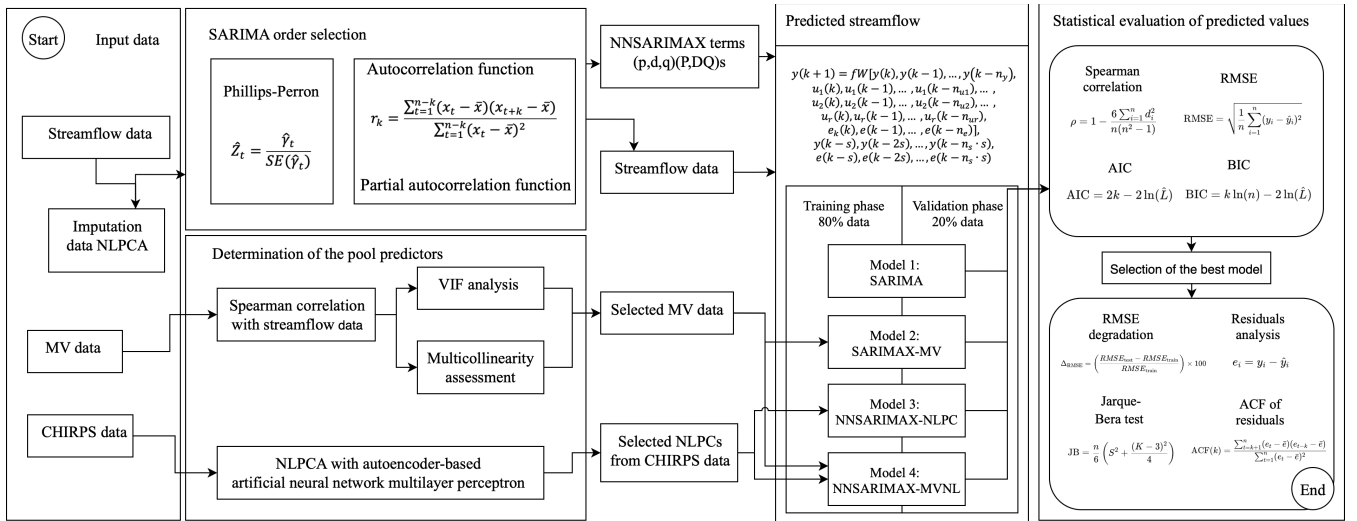

Figure 1: Schematic representation of the methodological workflow for streamflow forecasting under data-scarce conditions.

2.1 Study area

The Tocaría River originates at 3,200 m above sea level in the Guevarrica hill and drains a watershed of 2,223 km² on the eastern flank of the eastern cordillera. The river extends over 127.5 km with an average longitudinal slope of 5%. The basin is part of the Cravo Sur system, a tributary to the Meta River within the Orinoco macro-basin (Corporinoquia, 2010; Corporinoquia and Corpoboyacá, 2015b) (Fig. 2).

The basin experiences a monomodal rainfall regime with a distinct wet season from April to November and a dry season from December to March. Although 90% of the basin lies within the Andean region, its discharge contributes to the Orinoquía domain. Mean annual precipitation is approximately 2,031 mm, with peak streamflows during the wet season and minimum flows in February (Ruíz-Ochoa et al., 2022; Urrea et al., 2016).
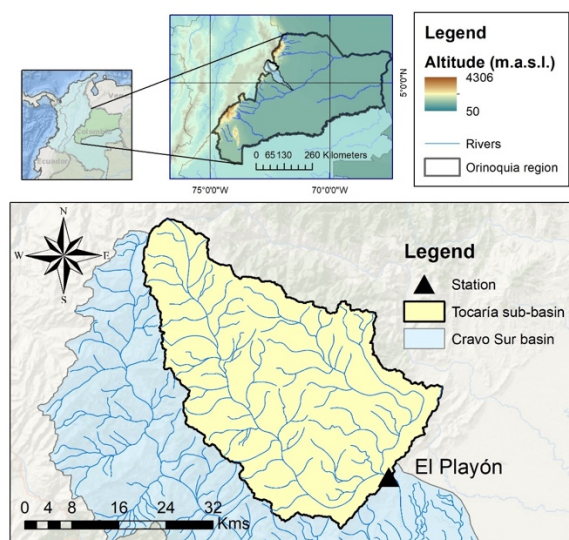
Figure 2: Geographic location of the Tocaría River basin and El Playón hydrometric station. Digital elevation model, political boundaries, watershed boundaries, and rivers extracted from GEOPORTAL IDEAM: https://archivo.ideam.gov.co/geoportal

2.2 Data

Three main datasets were used for model development: observed streamflow, satellite-based precipitation, and large-scale macroclimatic indices.

Streamflow data were obtained from the El Playón station, operated by Colombia's national hydrometeorological institute (IDEAM), covering 1983–2019 with approximately 5% missing values. Data were sourced from the DIHME geoportal (http://dhime.ideam.gov.co). Missing values were imputed using NLPCA, incorporating CHIRPS precipitation data as an exogenous input, following the approach by Ocampo-Marulanda et al. (2025).

Mean monthly discharge averaged 85 m³/s, with strong seasonal variability: February averaged 17 m³/s (driest month), while July peaked at 172 m³/s (Fig. 3). These fluctuations highlight the necessity of modeling strategies that capture both hydroclimatic seasonality and interannual variability.

4
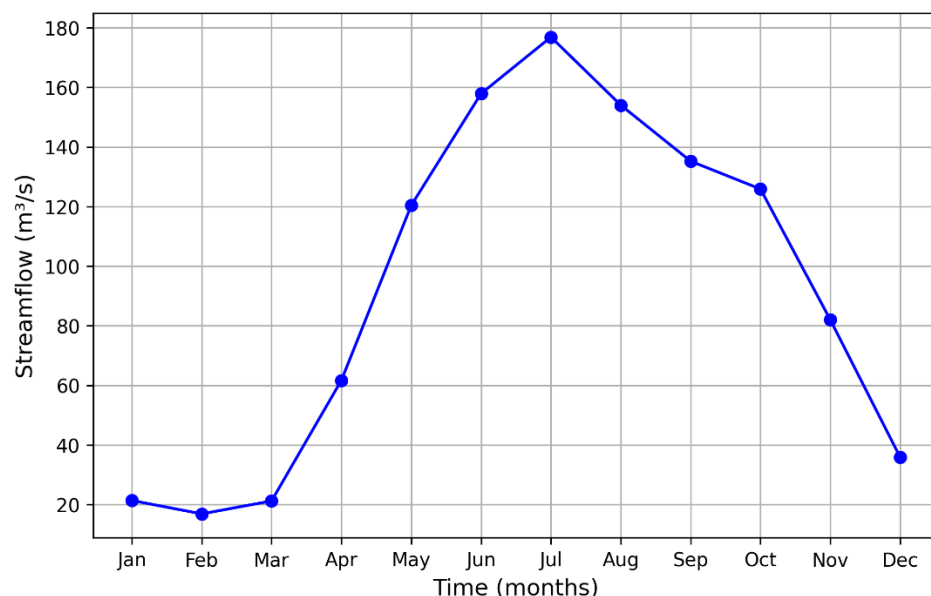
Figure 3: Monthly streamflow climatology at the El Playón station, Tocaría River.

100 Macroclimatic variables were retrieved from the National Oceanic and Atmospheric Administration (NOAA, https://psl.noaa.gov/gcos_wgsp/). Twenty-one indices representing sea surface temperature anomalies, atmospheric pressure, and wind patterns were selected based on prior demonstrated correlations with Colombian streamflow regimes (Poveda et al., 2011; Cerón et al., 2020; Canchala et al., 2020a). Complete descriptions of selected variables are available in Supplementary Table S1.

105 Precipitation data were sourced from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), which integrates satellite and in situ observations. CHIRPS provides global coverage at 0.05° spatial resolution from 1981 to present. Its performance in Colombia is well documented (Funk et al., 2014; Urrea et al., 2016; Ocampo-Marulanda et al., 2022). In this study, CHIRPS data were used to derive nonlinear principal components (NLPC) from 81 time series corresponding to 81 grid cells located within the Tocaría River basin.

110 **2.3 Data preprocessing**

This subsection details statistical preprocessing applied to streamflow data prior to model development, including variability characterization, trend detection, structural change identification, and stationarity assessment.

2.3.1 Descriptive statistics and trend analysis

Descriptive statistics calculated included mean (μ), maximum, standard deviation (σ), and coefficient of variation (CV), 115 providing an initial quantification of streamflow magnitude and variability.

Structural changes were identified via Pettitt's test (Pettitt, 1979), revealing a significant breakpoint in the time series mean. Sen's slope estimator (Sen, 1968) quantified long-term trends, while the Mann-Kendall test (Mann, 1945) evaluated trend significance.

2.3.2 Stationarity assessment

120 Stationarity was assessed through additive time series decomposition into trend, seasonal, and residual components, facilitating visualization of nonstationary behavior.

The Phillips–Perron (PP) test evaluated the null hypothesis of a unit root, offering robustness to autocorrelation and heteroskedasticity in residuals. Non-rejection of the null indicated nonstationarity, guiding NN-SARIMAX model structure selection.

125 ### 2.3.3 Autocorrelation and Partial Autocorrelation

Temporal dependencies were examined via autocorrelation function (ACF) and partial autocorrelation function (PACF) analyses (Chatfield, 1989). The ACF quantifies linear correlation at lag k, defined as:

$$r_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n-k}(x_t - \bar{x})^2} \tag{1}$$

where x_t is the series value at time t, $\bar{x}$ is the mean, and n the number of observations.

130 PACF measures correlation at lag k controlling for intermediate lags, informing autoregressive and moving average term selection in NN-SARIMAX models (Box et al., 2015). ACF and PACF plots were examined for seasonality and persistence, using 95% confidence intervals to identify significant lags.

2.4 Selection of possible correlators

2.4.1 CHIRPS precipitation as a predictor

135 The hydrological relationship between precipitation and streamflow in the Cravo Sur basin has been previously demonstrated (Ocampo-Marulanda et al., 2025), supporting the use of CHIRPS precipitation data as potential predictors. In this study, 81 precipitation time series corresponding to CHIRPS grid cells within the Tocaría sub-basin were analyzed.

To prevent overfitting and reduce dimensionality, NLPCA was applied to extract two principal components capturing dominant nonlinear spatiotemporal precipitation patterns. Implemented via an autoencoder with bottleneck architecture, NLPCA
140 minimizes reconstruction error between inputs and outputs, capturing nonlinear relationships beyond traditional PCA capabilities (Scholz et al., 2007; Canchala et al., 2019). Validation of extracted components included reconstruction accuracy and latent space visualization, confirming their efficacy in summarizing precipitation variability.

2.4.2 Macroclimatic variables as large-scale predictors

MVs comprising sea surface temperature anomalies, sea-level pressure, and wind indices known to influence Colombian

145 hydrology (Canchala et al., 2020a,b; Cerón et al., 2020; Poveda et al., 2001) were evaluated.

Spearman rank correlations between each MV and streamflow were calculated for lags up to 14 months to capture delayed effects. Variables exhibiting statistically significant correlations at lags ≥ 6 months were retained, reflecting hydroclimatic memory. Only predictors with demonstrated hydrological relevance and minimal intercorrelation were preserved to ensure model parsimony and stability.

2.4.3 Multicollinearity diagnosis and variable selection

150

To ensure stable and interpretable forecasting models, a two-step assessment of multicollinearity was performed. Initially, correlation matrices visualized as heatmaps identified groups of highly correlated predictors that could introduce redundancy and overfitting risks (Dormann et al., 2013). Subsequently, the Variance Inflation Factor (VIF) quantified the degree of linear dependence among predictors (Katrutska and Strijov, 2017). Predictors with VIF values exceeding the threshold of 10 were

155 flagged for potential exclusion.

However, exclusion was not automatic. Variables exhibiting high VIF but demonstrating strong and meaningful correlations with streamflow were retained to preserve model interpretability and relevance (Lavery et al., 2017). Conversely, predictors with both high VIF and weak correlation with the target variable were removed. This iterative process was coupled with evaluation of model performance across different predictor subsets, balancing parsimony, predictive accuracy, and explanatory

160 power. The final predictor set thus optimized robustness and transparency in streamflow forecasting.

2.5 Forecasting model: NN-SARIMAX architecture

The forecasting approach proposed herein is based on a hybrid NN-SARIMAX framework that integrates nonlinear temporal dynamics, exogenous climatic forcings, and seasonal variability within a unified predictive architecture. This model synergizes the statistical generalization capacity of SARIMA structures with the universal function approximation capabilities of neural

165 networks trained on NLPC extracted from high-dimensional hydroclimatic inputs.

The forecasted streamflow vector, denoted as $\hat{y}(k) \in R^m$, is modeled as a nonlinear function of multiple lagged sequences:

Streamflow autoregressive terms:

$$y(k), y(k-1), \dots, y(k-n_y) \tag{2}$$

Exogenous forcings (e.g., ocean-atmospheric indices):

170 $$u_1(k), u_1(k-1), \dots, u_1(k-n_{u1}), \dots; \dots; u_r(k), u_r(k-1), \dots, u_r(k-n_{ur}) \tag{3}$$

Forecast residuals:

$$e_k(k), e(k-1), \dots, e(k-n_e)] \tag{4}$$

Seasonal lags of streamflow and residuals:

$$y(k-s), y(k-2s), \dots, y(k-n_s \cdot s), \tag{5}$$

$$e(k-s), e(k-2s), \dots, e(k-n_s \cdot s) \tag{6}$$

The model's general transfer function is expressed as Ec(7):

$$y(k+1) = fW \begin{bmatrix} y(k), y(k-1), \dots, y(k-n_y); \\ u_1(k), u_1(k-1), \dots, u_1(k-n_{u1}); u_r(k), u_r(k-1), \dots, u_r(k-n_{ur}) \\ e_k(k), e(k-1), \dots, e(k-n_e)] \\ y(k-s), y(k-2s), \dots, y(k-n_s \cdot s), \\ e(k-s), e(k-2s), \dots, e(k-n_s \cdot s) \end{bmatrix} \tag{7}$$

Where n_y, n_{ur}, n_e, and n_s represent the autoregressive, exogenous, residual, and seasonal lag orders, respectively; s denotes seasonal periodicity; and fW is a nonlinear mapping function parameterized by model weights W.

This hybrid configuration enables simultaneous modeling of high-order dependencies, memory effects, and seasonal recurrence. Incorporating lagged residuals and seasonal components allows the model to capture patterns not accounted for by direct input–output relations, enhancing its capacity to learn both deterministic and stochastic dynamics.

An open-loop configuration was adopted during training, feeding observed streamflow values back to stabilize learning. For validation and operational forecasting, a closed-loop mode recursively propagates forecasted values. Residual sequences were standardized during validation and testing to ensure numerical stability.

The structure of the SARIMA model was determined through visual inspection of autocorrelation (ACF) and partial autocorrelation (PACF) plots of the monthly streamflow series. The orders (p,d,q) and (P,D,Q) were selected to reflect the presence of short-term dependencies and seasonal dynamics. Differencing was not applied, as the series appeared stationary in both its level and seasonal pattern. To support this assessment, the Phillips–Perron test was applied to evaluate the presence of unit roots and inform the choice of d=0 and D=0. A seasonal period of s=12 was defined based on the observed annual cycle in streamflow behavior.

Data were partitioned into 70% training, 15% validation, and 15% testing subsets, preserving temporal ordering to avoid information leakage; a blocked temporal split was preferred over random sampling. Alongside Bayesian regularization, early stopping was implemented by monitoring validation loss with a patience threshold of 100 epochs, halting training upon stagnation in performance.

Model performance was evaluated with complementary metrics reflecting accuracy and parsimony. RMSE quantified average prediction error magnitude and R^2 assessed explained variance. For NN-SARIMAX, model order selection criteria (AIC, BIC) balanced goodness-of-fit against model complexity.

2.6 Evaluation of the Models

2.6.1 Forecast Error Degradation Across Prediction Horizons

Forecasting accuracy typically decreases as the prediction horizon extends, a well-documented phenomenon referred to as forecast error degradation (Xiang et al., 2024). To assess this behavior, the best-performing NN-SARIMAX model was evaluated over multiple lead times using RMSE and R^2 as performance metrics. Forecasts were produced for 2 to 24 months ahead in increments of two months, enabling the quantification of predictive accuracy across short, medium, and long term horizons.

To capture the evolution of forecast quality, RMSE and R^2 values were analyzed in pairs of consecutive horizons. This pairing facilitated the visualization of degradation trends through line plots, offering insights into the temporal limits of model effectiveness. The resulting performance curves contribute to a clearer understanding of how predictive uncertainty accumulates over time, which is critical for operational decision-making in hydrological forecasting.

2.6.2 Residual Analysis

Residual diagnostics are essential to evaluate model adequacy by verifying whether prediction errors conform to standard statistical assumptions. Residuals $e_t = y_t - \hat{y}_t$ were computed as the difference between observed and forecasted streamflow values, forming a time series subjected to rigorous statistical and graphical analysis.

Normality was tested using the Jarque–Bera test (Jarque and Bera, 1980), assessing whether residuals follow a Gaussian distribution, a prerequisite for valid inference and uncertainty quantification. Autocorrelation was examined through the Durbin–Watson statistic (Durbin and Watson, 1950), where values near 2 suggest the absence of significant serial dependence, thus confirming that temporal structures have been effectively captured. Heteroscedasticity was assessed via the Breusch–Pagan test (Breusch and Pagan, 1979); significant p-values indicated non-constant residual variance, potentially undermining the model's generalizability.

Complementing these tests, graphical diagnostics were employed to detect patterns indicative of model misspecification. Histograms with kernel density estimation (Silverman, 1986) visually assessed normality, while residuals versus fitted values plots revealed heteroscedastic patterns or structural biases. Quantile–quantile plots (Wilk and Gnanadesikan, 1968) further tested alignment between empirical and theoretical residual quantiles. Finally, ACF plots identified lagged dependencies, where the absence of significant peaks outside confidence bounds supported the white noise assumption.

3 RESULTS AND DISCUSSIONS

The experimental results are organized into three subsections: (1) Description of the streamflow data, (2) Determination of the pool predictors, and (3) Forecasted streamflow series. Each subsection presents and discusses relevant findings, emphasizing the identification of the most suitable forecasting model based on accuracy metrics. A comparative analysis of the four

evaluated models highlights their strengths and limitations, guiding the selection of the most effective model for streamflow
230   prediction.

Long-range streamflow forecasting is a valuable tool for water resource managers, enabling proactive decision-making and
strategic planning. However, excessively extending the forecast lead-time tends to degrade model performance, often resulting
in diminished predictive accuracy. Consequently, this study adopts a lead-time of 24 months to evaluate and compare
forecasting models. This horizon provides a balance between planning requirements and achievable accuracy for watershed
235   management stakeholders.

3.1 Description of the Streamflow Data

3.1.1 Statistical and Graphical Description

The streamflow time series at El Playón station exhibits pronounced interannual and seasonal variability, with values ranging
from 2.5 m³/s to 280 m³/s and a mean of 92.4 m³/s. This amplitude is consistent with the monomodal rainfall regime
240   characteristic of the region (Urrea et al., 2019). A standard deviation of 66.4 m³/s and a coefficient of variation of 71.8%
confirm the high dispersion and temporal heterogeneity of the data, which poses a challenge for accurate modeling and
prediction.

Trend detection tests were also applied. A visual inspection suggested a potential breakpoint in January 2019; however, the
Pettitt test yielded a p-value of 1.99, indicating that this change is not statistically significant. Likewise, Sen's slope and the
245   Mann-Kendall test indicate a slight upward trend in discharge, but without statistical significance (p = 0.3445), suggesting the
absence of a persistent monotonic trend (see Table 1 and Fig. 4).

Table 1: Basic statistics and trend characterization for El Playón station flow rates.

| Mean	Min	Max	Standard deviation	Coefficient of variation	Break poin	Pettitt (p-value)	Sen slope	Mann-Kendall (p-value)
92.4	2.5	280	66.4	71.8	Jan/2019	1.99	0.0188	0.3445

250
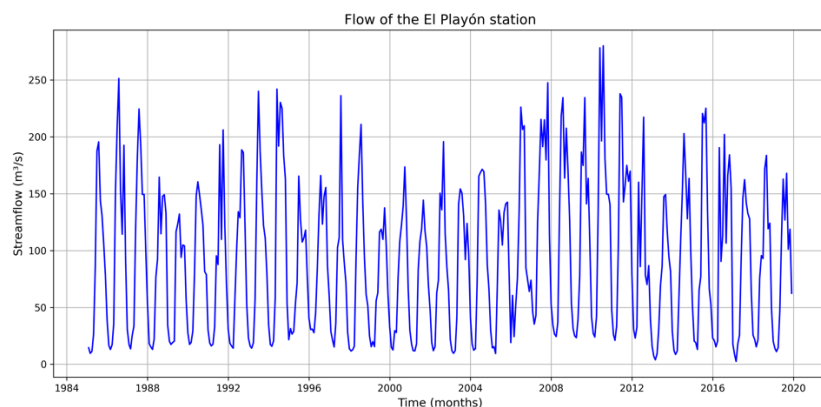
Figure 4: Temporal behavior of streamflow at the Tocaría River.

These characteristics highlight the complexity of the time series and justify the need for rigorous preprocessing, including stationarity and autocorrelation analysis.

255   **3.1.2 Stationarity Analysis**

Stationarity was assessed through both graphical decomposition and formal statistical testing. The decomposition revealed a marked seasonal component and residuals without systematic trend or variance shifts, supporting weak stationarity assumptions component lacking any systematic trend, indicating time-invariance in mean and variance (Figure 5).
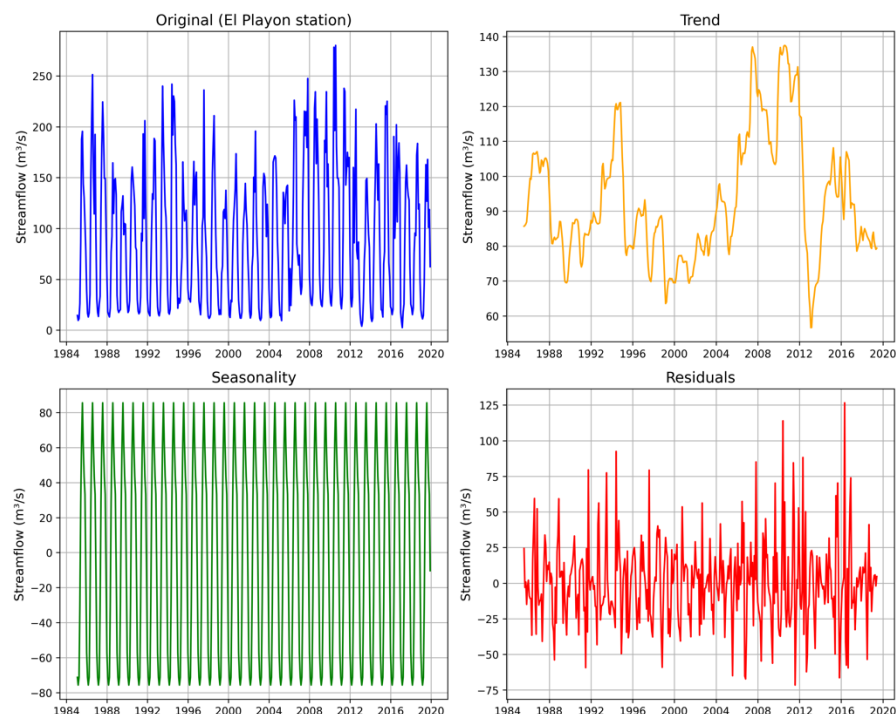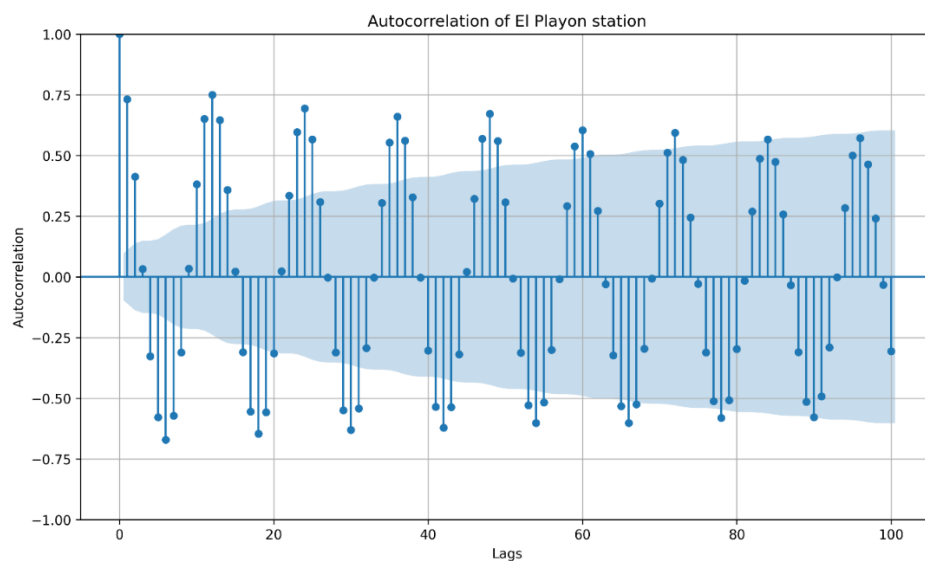
Figure 5: Decomposition of El Playón streamflow time series into trend, seasonality, and residual components.

The Phillips-Perron test further confirmed stationarity, yielding a test statistic of -6.5005 and a p-value < 0.0001, allowing rejection of the null hypothesis of a unit root. These results justify modeling without differencing and the selection of $d = 0$ in the SARIMA model.
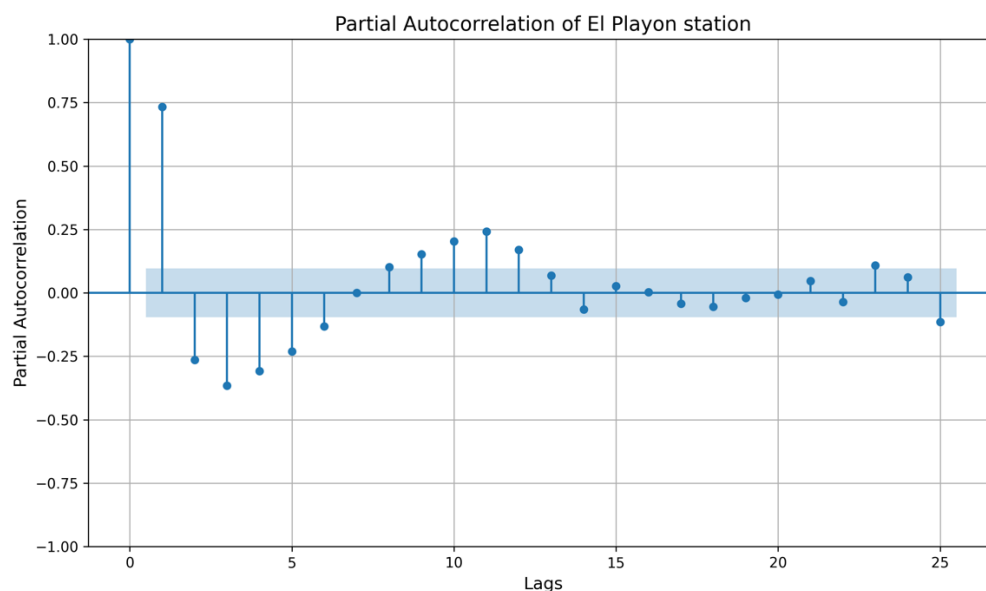
3.1.3 Autocorrelation and Partial Autocorrelation Structure

The ACF displays significant periodic peaks at regular intervals, indicative of strong seasonal periodicity and persistent temporal dependencies beyond short lags (Fig. 6). This cyclicality aligns with hydrological patterns driven by seasonal precipitation and runoff dynamics, underscoring the appropriateness of seasonal forecasting models such as SARIMA.

Figure 6: Autocorrelation of the time series of El Playón streamflow for 100 months of lags.

The PACF shows a prominent spike at lag 1 followed by rapid decay to values within the confidence bounds (Fig. 7), consistent with an autoregressive process of order 1 (AR(1)) though the selected SARIMA model incorporates higher-order dynamics to capture the longer memory and seasonal structure identified in the ACF. The lack of significant PACF values at higher lags indicates limited benefit from additional autoregressive terms.

Figure 7: Partial autocorrelation function of El Playón streamflow series over 25 monthly lags.

13

280

Based on combined ACF and PACF analysis, a SARIMA $(4,0,4)(0,0,3)_{12}$ model was selected. This model configuration incorporates the identified seasonal periodicity (season length $s = 12$), confirms stationarity ($d = 0$), and captures short-term autoregressive and moving average dynamics. Residual diagnostics validated the model's adequacy, demonstrating no significant autocorrelation.

3.2. Determination of the Pool Predictors

3.2.1 Selection of precipitation NLPCs as exogenous variables

The 81 CHIRPS precipitation time series for the Tocaría River basin were reduced to two nonlinear principal components (NLPC1 and NLPC2) using NLPCA. NLPC1 explains 92.5% of the total variance (eigenvalue = 25.6), and NLPC2 accounts for 7.4% (eigenvalue = 2.1), preserving the dominant spatiotemporal patterns of the original dataset retaining over 99% of the total variance and the key spatiotemporal signatures of basin-wide precipitation.

Spearman correlation analysis revealed moderate positive correlations between the NLPCs and the streamflow observed at the El Playón station, with coefficients of 0.50 for NLPC1 and 0.34 for NLPC2. These results support the predictive utility of the NLPCs as exogenous variables, consistent with findings by Ocampo-Marulanda et al. (2025), who used CHIRPS-derived precipitation components to reconstruct streamflow time series in the Cravo Sur basin under data-scarce conditions. These components were subsequently included as covariates in the streamflow forecasting model, enabling a more parsimonious yet informative representation of the basin's precipitation dynamics

Incorporating NLPCs as exogenous predictors offers a compact yet informative representation of precipitation, reducing dimensionality without sacrificing critical hydrometeorological signals.

3.2.2 Selection of Macroclimatic Variables as Exogenous Variables

The influence of 21 MVs on the streamflow variability of the Tocaría River was assessed using a lagged Spearman correlation analysis, considering time lags from 0 to 14 months. This asynchronous framework allowed for identification of delayed relationships between climate anomalies and hydrological responses, which are crucial for predictive modeling. This lagged correlation framework enables detection of delayed hydroclimatic teleconnections, which are essential for long-lead streamflow forecasting.

Statistical significance was evaluated using two-tailed tests with confidence levels between 90% and 99%. Figure 8 displays the lagged correlations between the MVs and the streamflow at the El Playón station. Several variables demonstrated statistically significant correlations across multiple lags, exhibiting a quasi-periodic structure that reflects the hydroclimatic seasonality modulated by the Intertropical Convergence Zone (ITCZ).

14
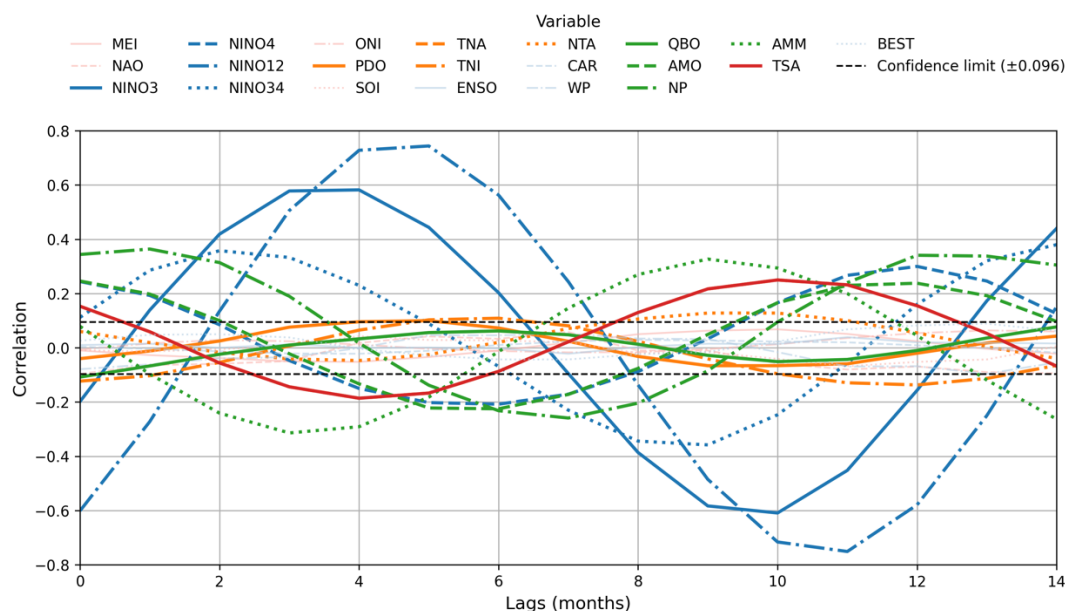
Figure 8: Lagged Spearman correlations between MVs and streamflow at El Playón station.

Key Pacific-related indices—namely NINO12 ($\rho = 0.75$) and NINO3 ($\rho = 0.60$)—emerged as the strongest predictors, underscoring the robust influence of ENSO-related sea surface temperature (SST) anomalies on the Tocaría River streamflow. These findings are consistent with previous studies across Colombia (e.g., Canchala, 2020c; Cadavid and Salazar, 2008; Poveda et al., 2002), which reported comparable lagged correlations between ENSO indices and regional streamflow, often with cyclic behavior.

In addition, Atlantic-related variables such as AMM ($\rho = 0.32$) and TSA ($\rho = 0.25$) showed statistically significant, albeit weaker, correlations. These results corroborate the relevance of Atlantic SST variability in eastern Colombia. The Llanos low-level jet and moisture transport from the tropical Atlantic—modulated by indices such as AMM and TSA—have been highlighted as key controls of precipitation and streamflow in the Orinoquía region (Labat et al., 2012; Nieto et al., 2008; Builes et al., 2022; Correa et al., 2024).

The set of significantly correlated predictors spans both Pacific and Atlantic basins: NINO3, NINO4, NINO12, NINO34, PDO, TNI, NP, QBO (Pacific-related), and AMO, NTA, TNA, NAO, AMM, TSA (Atlantic-related). This reinforces the complexity of the climate–hydrology interactions in the Tocaría River basin and emphasizes the need to account for both direct and lagged macroclimatic signals in predictive modeling.

3.2.3 Assessing Multicollinearity and Variable Selection for Improved Time Series Forecasting

To ensure model robustness and avoid overfitting, multicollinearity among predictors was quantified using the VIF. Table 2 presents the VIF values for all candidate variables. While most variables showed acceptable levels (VIF < 10), certain

330 predictors; especially those related to ENSO; exhibited extreme multicollinearity. NINO3, NINO4, NINO12, and NINO34 had VIFs ranging from 2,492 to 81,825. TNA and TSA presented infinite VIFs, indicating perfect collinearity with other variables most likely due to structural redundancy among Atlantic SST indices, as confirmed by the correlation heatmap.

Table 2: Variance Inflation Factor Analysis for predictor variables.

Variable	VIF	Variable	VIF
NAO	2	NTA	16
NINO3	38,899	QBO	1
NINO4	29,126	AMO	8
NINO12	2,492	NP	17,705
NINO34	81,825	AMM	8
PDO	2	TSA	inf
TNA	inf	PNLPCA_PC1	4
TNI	4	PNLPCA_PC2	4

335

These results, along with the correlation heatmap (Fig. 9), confirm substantial interdependence among ENSO indices, suggesting that including multiple related predictors may lead to instability and redundancy. For example, NINO3, NINO4, and NINO34 are highly correlated, and thus only one may be sufficient to represent ENSO dynamics. Likewise, high correlation between TNA and TNI implies redundancy.

340
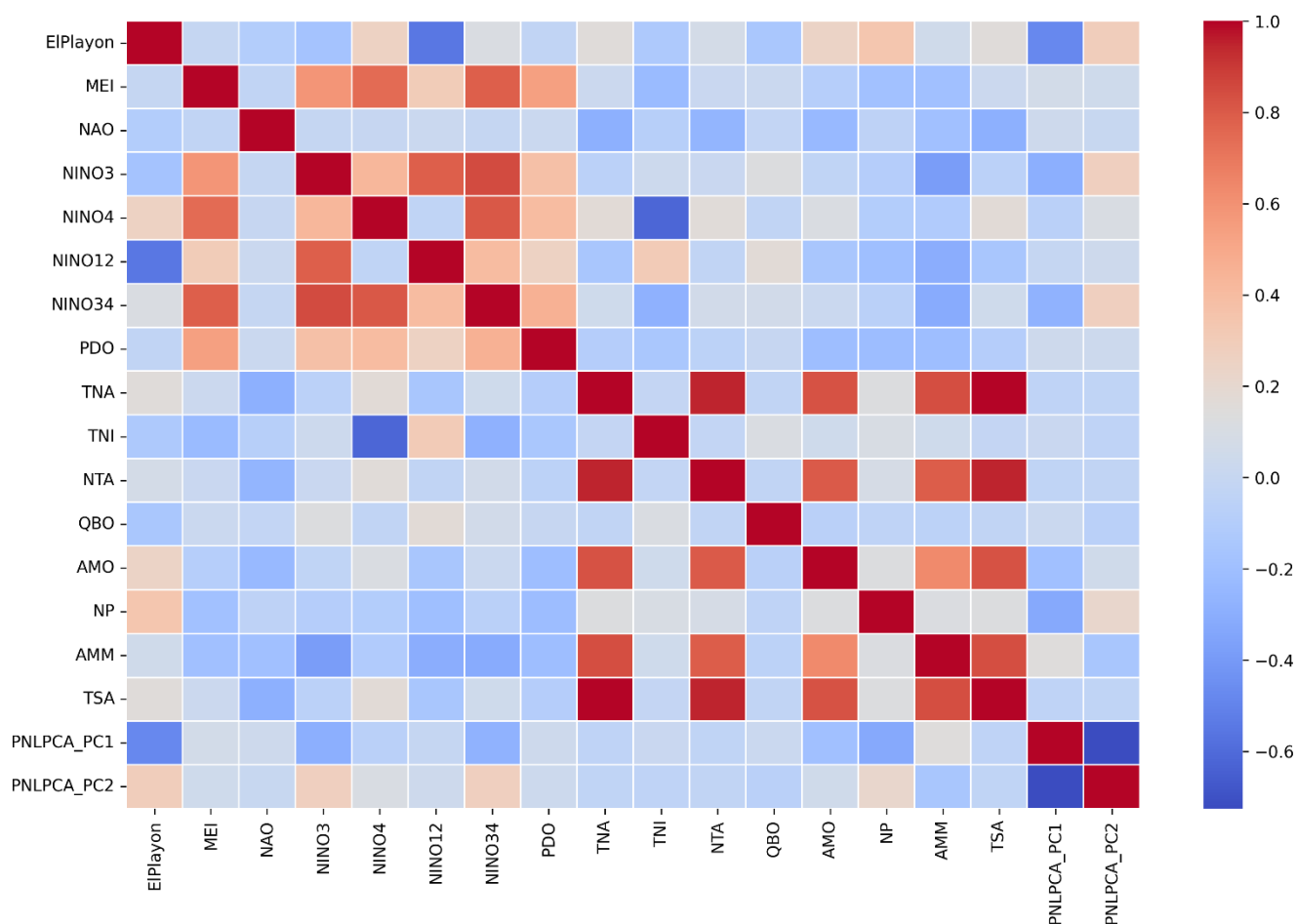
Figure 9: Correlation heatmap of predictor variables.

To address this, various combinations of variables were tested through an iterative model refinement process, prioritizing

345 predictors with strong individual relationships to streamflow and minimal mutual collinearity. This selection strategy balances predictive power with model parsimony, mitigating multicollinearity-induced variance inflation and enhancing generalizability. Model performance was evaluated using RMSE and R^2 to balance predictive accuracy and parsimony.

Ultimately, the optimal subset of exogenous predictors was determined to be: PNLPCA_PC1, PNLPCA_PC2, NINO4, NINO12, NP, AMO, TNA, AMM, and TSA.

350 This final configuration minimized multicollinearity, preserved critical information from both precipitation and climate variability, and improved the interpretability and reliability of the streamflow forecasting model.

The final pool of predictors; comprising key macroclimatic indices and the first two NLPCs; provided the input for the forecasting models evaluated in the subsequent section. The next subsection presents the performance comparison of candidate models and the resulting streamflow forecasts.

3.3 Forecasted Streamflow Series

The forecasting performance of the four models was evaluated using R^2, RMSE, AIC, and BIC. All models were trained and validated on independent datasets to ensure generalization. Table 3 summarizes the results. All configurations yielded adequate performance in forecasting monthly streamflow up to 24 months in advance. As expected, R^2 values decreased and RMSE increased from training to validation. Training R^2 ranged from 0.85 to 0.91, while validation R^2 dropped to 0.71–0.78. RMSE increased from 26.8–32.7 in training to 41.5–49.9 in validation.

The model combining ocean-atmospheric indices with nonlinear components showed the smallest degradation between training and validation, indicating higher robustness. In contrast, the baseline and the model using only nonlinear components yielded lower predictive skill, though both maintained acceptable performance (validation R^2 of 0.72 and 0.71, respectively), which is notable given the hydroclimatic variability of the basin.

The best results were obtained with the combined model, which achieved the highest R^2 (0.91 training, 0.78 validation) and the lowest RMSE (26.8 and 41.5). No other model exceeded 0.87 in training or 0.73 in validation. All other models reported training RMSE above 29.0 and validation RMSE above 44.4. The combined configuration consistently showed lower errors.

In terms of model complexity, AIC was identical (3103) for the three simpler configurations, while slightly higher (3110) in the combined model. BIC increased more noticeably in the combined case (3173), reflecting its additional structural complexity.

Although the inclusion of nonlinear components alone did not significantly enhance the baseline, their integration with ocean-atmospheric variables resulted in a marked improvement in accuracy and consistency across datasets.

Table 3: Forecast performance metrics for the four streamflow forecasting models.

Model	Exogenous variables (inputs)	Training		Validation		AIC	BIC
		R^2	RMSE	R^2	RMSE		
SARIMA (4,0,4) (0,0,3)$_{12}$	None	0.86	30.2	0.72	49.9	3103	3122
NNSARIMAX-MV	MV	0.87	29.0	0.73	44.4	3103	3129
NNSARIMAX-NLPC	NLPC	0.85	32.7	0.71	49.7	3103	3129
NNSARIMAX-MVNL	MV and NLPC	0.91	26.8	0.78	41.5	3110	3173

These results are consistent with previous findings, Table 4, in the literature. Reported R² values for monthly streamflow prediction range from moderate (~0.65) to near-perfect (~0.99), depending on lead time, predictor variables, and catchment complexity.

For instance, Hosseinzadeh et al. (2023) reported R² = 0.92 using a multivariate SARIMAX model with temperature and
380 precipitation as exogenous inputs in the Colorado River Basin. Cheng et al. (2020) achieved R² = 0.95 using an LSTM model, although this performance was limited to one-month-ahead forecasts and dropped significantly (R² < 0.20) at longer lead times. Several hybrid or AI-based approaches have yielded high predictive accuracy (R² ≈ 0.99), such as those reported by Fathian (2019), Moeeni et al. (2017), and Dariane (2024). However, these results often stem from short-term horizons, multi-source calibration, or highly tuned model configurations prone to overfitting. Ghorbani (2016) obtained R² between 0.77 and 0.84
385 using ANN and SVM models, while Gómez (2010) reported a more moderate R² ≈ 0.64 for the Bogotá River basin.

In comparison, the performance of our models; particularly the NN-SARIMAX-MVNL; stands competitively within this broader context. Despite relying on a single hydrometric station and forecasting beyond the one-month horizon, our approach demonstrates robust skill in capturing the high temporal variability and structural complexity of streamflow in the Tocaría River Basin.

390

Table 4: Validation R² of selected monthly streamflow forecasting models from literature.

Best model	Description	Validation R²	References
SARIMAX	24-month forecasts with temperature and precipitation	0.92	Hosseinzadeh, et al., 2023
LSTM	One-month-ahead forecasts	0.95	Cheng et al., 2020
MARS1-SETAR	Hybrid nonlinear time series and AI modeling	0.99	Fathian, et al 2019
SARIMA-ANFIS	SARIMA combined with neuro-fuzzy and ANN systems	0.94	Moeeni, 2017
Multistep data-driven	Neural-fuzzy hybrid with input selection	0.97	Dariane, 2017
Hybrid model	SARIMA combined with nonlinear ANN	0.72	Moeeni, 2016
ELM	Extreme Learning Machine approach	0.82	Yassen et al., 2016
RBF	MLP, RBF, and SVM models for river flow prediction	0.87	Ghorbani, 2016
PMC 20 NCO	Neuro-fuzzy vs neural networks for Bogotá River	0.65	Gomez, 2012

3.3.1 Statistical Evaluation of the Best Model for Streamflow Forecasting

Figure 10 shows the streamflow prediction results for the El Playón station using the NN-SARIMAX-MVNL model, which integrates exogenous variables including MV's and NLPCs derived from local precipitation data. The figure includes three key elements: observed streamflow, model predictions for the training period, forecasts for the validation period, and test period.

The observed streamflow series reveals pronounced seasonality, characterized by recurrent peaks and troughs consistent with the hydroclimatic regime of the basin. During the training phase, the model reproduces the seasonal dynamics and intra-annual variability with high accuracy, showing minimal deviations. In the validation period, although some differences appear; particularly underestimations of extreme highs or delays in response; the general flow patterns, including the timing and intensity of seasonal peaks, are well preserved. These results highlight the model's capacity to capture both historical and prospective streamflow behaviors based on relevant exogenous predictors.
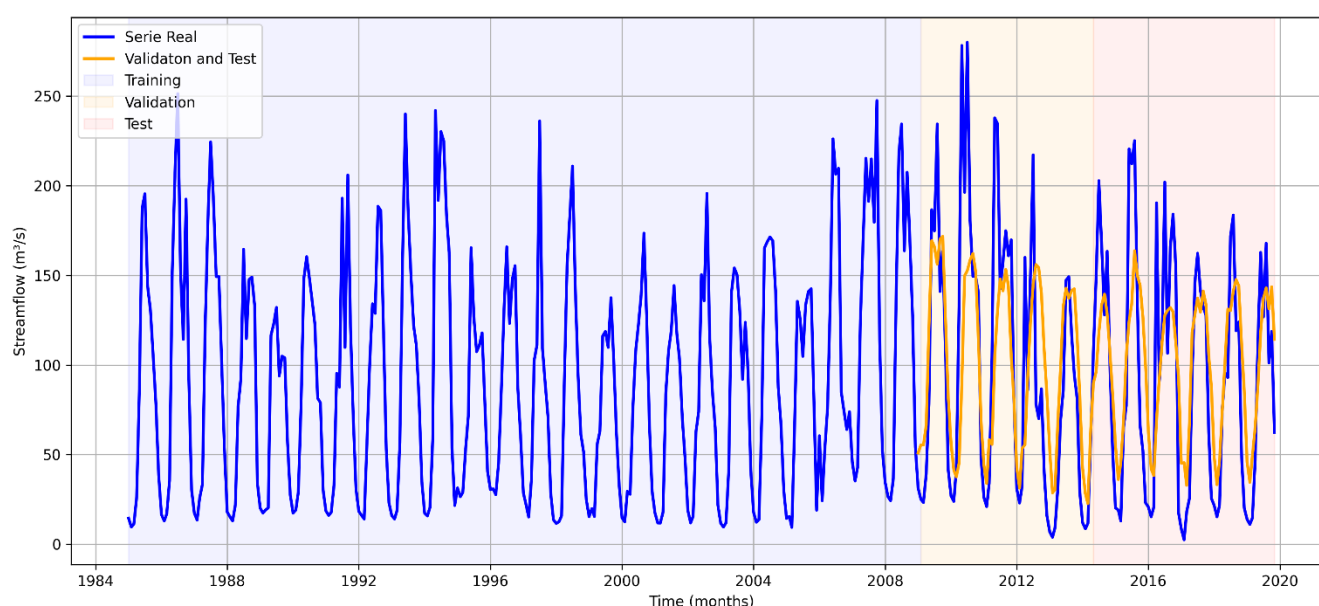

Figure 10: Observed and predicted streamflow for training and validation using the NN-SARIMAX-MVNL model.

To evaluate the model's temporal generalization capability, streamflow forecasts were generated for horizons up to 24 months (Fig. 11). The RMSE remained below 11 m³/s for the first four months, which is low relative to the standard deviation of the observed series (σ ≈ 66 m³/s). Between months 5 and 8, RMSE increased to approximately 30–50 m³/s and then stabilized. The R^2 metric remained above 0.90 up to month 7, then gradually declined, reaching 0.83 by month 10 and remaining above 0.75 throughout the entire 24-month forecast horizon.

20

These results confirm the robustness of the NN-SARIMAX-MVNL model for medium- and long-term forecasting, maintaining

415 a favorable balance between accuracy and lead time. The performance is suitable for practical decision-making applications in water resource planning, including drought early warning and integrated watershed management in the Tocaría River basin.
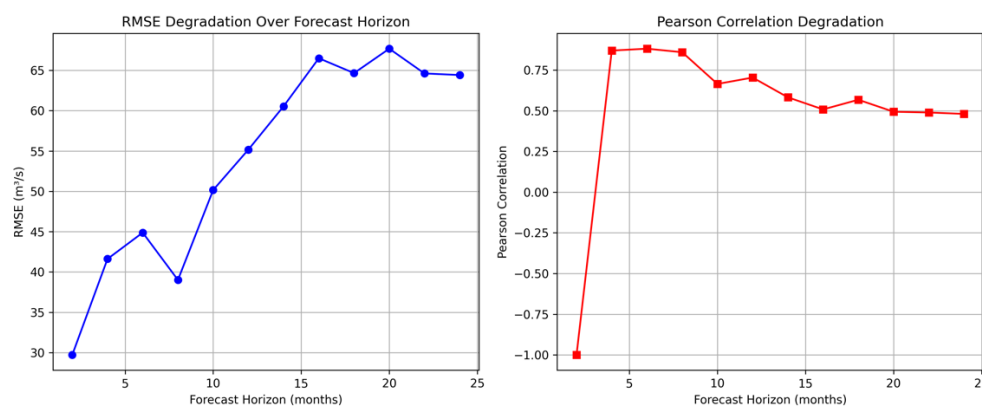
Figure 11: Forecast error degradation across prediction horizons based on RMSE and R² metrics.

420

A residual analysis was performed to assess the adequacy of the model and detect any remaining structure in the errors. Residuals were computed as the difference between observed and predicted streamflows. The residual series is generally centered around zero, with sporadic peaks reaching +150 m³/s and troughs of −100 m³/s, primarily associated with abrupt hydrological events.

425 Several statistical diagnostics were applied to evaluate the residual distribution. The Jarque–Bera test yielded a statistic of 13.2 (p = 0.0014), suggesting a moderate deviation from normality. The Q–Q plot reveals mild departures from the theoretical normal line, particularly in the upper quantiles, indicating that extreme residuals are somewhat asymmetrically distributed (Fig. 12). Overall, while the residuals are not perfectly normal, their distribution remains reasonably close to symmetry for most values.

430 The Durbin–Watson statistic of 1.72 indicates an absence of significant autocorrelation, which is corroborated by the ACF plot, where no substantial lags exceed the confidence bounds. This implies that the model has adequately captured the temporal dependencies in the streamflow series, and that residuals are largely uncorrelated.

The Breusch–Pagan test returned a statistic of 19.05 (p = 0.0396), suggesting the presence of heteroscedasticity. This is further supported by the residuals vs. fitted values plot, where residual variance appears non-constant, particularly for low to mid-

435 range predicted values. Clustering and trends in residuals are visible in the initial portion of the fitted value range, consistent with heteroscedastic behavior. These findings indicate that while the model performs well overall, its variance may vary across the prediction space, particularly during high variability periods.
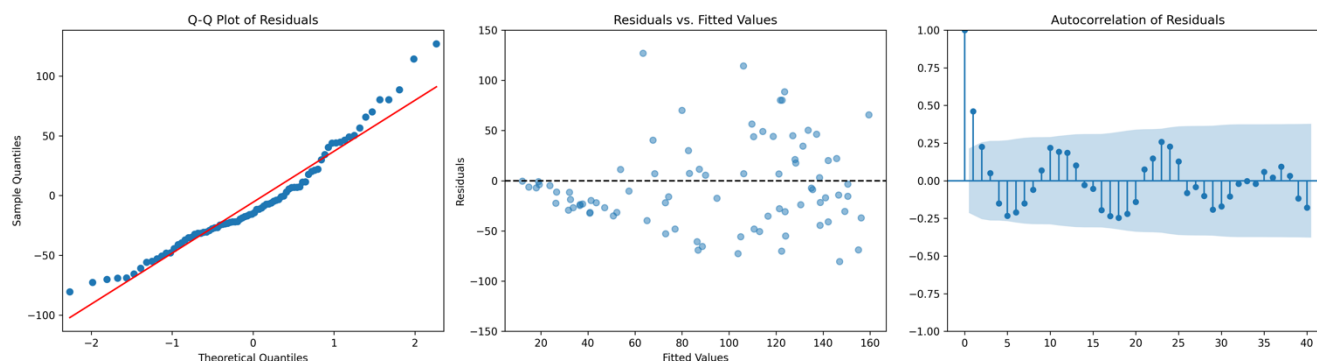
Figure 12: Graphical analysis of normality (Q-Q-Plot of residuals), heteroscedasticity (scatter plot, Residuals vs Fitted values), and ACF of residuals.

In summary, the statistical analysis supports the conclusion that the NN-SARIMAX-MVNL model provides reliable streamflow forecasts under average and low-flow conditions. However, its ability to accurately predict high-magnitude peak flows remains limited, likely due to complex hydrometeorological drivers not fully captured by the selected predictors. Despite this, the model demonstrates strong potential for operational use in drought risk assessment and water resource management, particularly for forecasting low-flow events in the Tocaría River basin.

4 CONCLUSIONS

This study conducted a rigorous comparative assessment of four monthly streamflow forecasting models for a data-scarce basin, all based on the SARIMA framework. These models were incrementally enhanced through the integration of exogenous predictors and ANNs to improve predictive skill. All configurations demonstrated satisfactory performance, as reflected in high R^2 and low RMSE across both calibration and validation phases. However, the baseline SARIMA $(4,0,4)(0,0,3)_{12}$ model, while capable of capturing seasonal and autoregressive patterns, showed limited forecasting ability when compared to models incorporating external predictors.

Among the configurations tested, the SARIMAX models with MVs, and the combination of MVs with NLPCs; substantially outperformed the baseline. The NN-SARIMAX-MVNL model, which merges both large-scale climate signals and localized precipitation variability, achieved the highest forecasting accuracy and stability, consistently yielding the lowest RMSE and highest R^2 in both training and out-of-sample testing.

The selection of exogenous predictors was grounded in hydroclimatic relevance. The MVs included indices such as NINO4, NINO12, NP, AMO, TNA, AMM, and TSA, which collectively capture dominant modes of Pacific and Atlantic ocean–atmosphere variability affecting the Tocaría River basin. At the local scale, precipitation data derived from the CHIRPS dataset

were incorporated via NLPCA, allowing the model to account for spatially distributed and nonlinear hydrometeorological influences.

465   The NN-SARIMAX-MVNL model demonstrated strong predictive performance at extended lead times of up to 24 months, highlighting its applicability for anticipatory water resource management and hydroclimatic risk mitigation in data-limited settings. Residual diagnostics confirmed its reliability in forecasting both low; and average-flow regimes; an essential feature for drought preparedness and long-term planning in the basin.

In summary, this work underscores the value of integrating large-scale climate teleconnections with locally resolved
470   precipitation dynamics to enhance streamflow predictability under conditions of data scarcity. The hybrid modeling framework proposed herein offers a scalable and transferable approach for streamflow forecasting in other poorly instrumented basins, with implications for strengthening adaptive watershed governance in the context of climate variability and change.

5 AUTHOR CONTRIBUTION

Conceptualization: JDSO, COM, LMCA, TC, and TAFE; data curation: JDSO and COM; formal analysis: JDSO, COM;
475   LMCA, TC, and TAFE; Funding acquisition: LMCA; Methodology: COM, TC, and TAFE; Project administration: LMCA; Resources: LMCA and TAFE; Software: JDSO, COM; Supervision: LMCA, TC and TAFE; Validation: TC and TAFE; Visualization: JDSO and COM; Writing (original draft preparation): JDSO, COM, LMCA and TC; and Writing (review and editing): JDSO, COM; LMCA, TC, and TAFE.

6 COMPETING INTERESTS

480   The authors also have no other competing interests to declare.

7 ACKNOWLEDGEMENTS

8 FINANCIAL SUPPORT

9 REFERENCES

Builes-Jaramillo, A., Yepes, J., Salas, H.D.: The Orinoco low-level jet and its association with the hydroclimatology of northern South America, J. Hydrometeorol., 23, 209–223, https://doi.org/10.1175/JHM-D-21-0073.1, 2022.

495   Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: Time Series Analysis: Forecasting and Control, 5 ed., John Wiley and Sons, 2015.

Breusch, T.S., Pagan, A.R.: A Simple Test for Heteroscedasticity and Random Coefficient Variation, Econometrica, 47, 1287-1294, http://dx.doi.org/10.2307/1911963, 1979.

Cadavid, J.M., Salazar, J.E.: Generación de series sintéticas de caudales usando un modelo Matalas con medias condicionadas,

500   Av. Recur. Hídricos, 17, 2008.

Canchala-Nastar, T., Carvajal-Escobar, Y., Alfonso-Morales, W., and Caicedo-Bravo, E.: Estimation of missing data of monthly rainfall in Southwestern Colombia using artificial neural networks. Data in Brief, 26, 104517, https://doi.org/10.1016/j.dib.2019.104517, 2019.

Canchala, T., Cerón, W. L., Frances, F., Carvajal-Escobar, Y., Andreoli, R. V., Kayano, M. T., Alfonso-Morales, W.,

505   Caicedo-Bravo, E., Ferreira de Souza, R. A.: Streamflow Variability in Colombian Pacific Basins and Their Teleconnections with Climate Indices. Water, 12, 526, https://doi.org/10.3390/w12020526, 2020a.

Canchala, T., Alfonso-Morales, W., Cerón, W. L., Carvajal-Escobar, Y., Caicedo-Bravo, E.: Teleconnections between Monthly Rainfall Variability and Large-Scale Climate Indices in Southwestern Colombia. Water, 12, 1863, https://doi.org/10.3390/w12071863, 2020b.

510   Canchala, T., Alfonso-Morales, W., Carvajal-Escobar, Y., Cerón, W. L., Caicedo-Bravo, E.: Monthly Rainfall Anomalies Forecasting for Southwestern Colombia Using Artificial Neural Networks Approaches. Water, 12, 2628, https://doi.org/10.3390/w12092628, 2020c.

Cárdenas-Rodríguez, S., Vides-Herrera, A., Pardo-García, A.: Sistema de alerta temprana de inundaciones para el río Arauca basado en técnicas de inteligencia artificial, Rev. Investig. Desarr. Innov., 12(2), 315–326, 2022.

515   Cerón, W. L., Andreoli, R. V., Kayano, M. T., Ferreira de Souza, R. A., Jones, C., Carvalho, L. M. V.: The Influence of the Atlantic Multidecadal Oscillation on the Choco Low-Level Jet and Precipitation in Colombia. Atmosphere, 11, 174, https://doi.org/10.3390/atmos11020174, 2020.

Chatfield, C.: The Analysis of Time Series: An Introduction, 4 ed., Chapman and Hall, 1989.

Cheng, M., Fang, F., Kinouchi, T., Navon, I.M, Pain, C.C.: Long lead-time daily and monthly streamflow forecasting using

520   machine learning methods. Journal of Hydrology, 590, https://doi.org/10.1016j.jhydrol.2020.125376. 2020.

Chiew, F.H.S., McMahon, T.A.: Teleconnections between the ENSO phenomenon and streamflow, streamflow forecasting, and interannual variability, Hydrol. Sci. J., 47(3), 505–522, https://doi.org/10.1080/02626660209492950, 2002.

Corporinoquia: Resolución No. 200.41-10.1402 — Por medio de la cual se regula el uso y aprovechamiento del recurso hídrico en el río Tocaría, https://www.corporinoquia.gov.co (últ. acc. 20 jun 2025), 2010.

525 Corporinoquia, Corpoboyacá: Evaluación de la zonificación de las amenazas en la cuenca del río Cravo Sur (Cap. 6), en: Plan de Ordenación y Manejo de la Cuenca del Río Cravo Sur – POMCA, 1–20, 2015a.

Corporinoquia, Corpoboyacá: Componente hidrológico (Cap. 3), en: Plan de Ordenación y Manejo de la Cuenca del Río Cravo Sur – POMCA, 1–22, 2015b.

Correa, I., Arias, P., Vieira, S., Martínez, A.: A drier Orinoco basin during the twenty-first century: The role of the Orinoco
530 low-level jet, Clim. Dyn., 62, 2369–2390, https://doi.org/10.1007/s00382-023-07028-7, 2024.

Costa, G.E.D.M.E., Menezes Filho, F.C.M.D., Canales, F.A., Fava, M.C., Brandão, A.R.A., de Paes, R.P.: Assessment of time-series models for mean discharge modeling and forecasting in a sub-basin of the Paranaíba River, Brazil, Hydrology, 10(11), 208, https://doi.org/10.3390/hydrology10110208, 2023.

Dariane, A.B., Behbahani, M.M.: Maximum energy entropy: A novel signal preprocessing approach for data-driven monthly
535 streamflow forecasting, Ecol. Inform., 79, 102452, https://doi.org/10.1016/j.ecoinf.2023.102452, 2024.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Münkemüller, T.: Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36, 27–46, https://doi.org/10.1111/j.1600-0587.2012.07348.x, 2013.

Durbin, J.; Watson, G. S.:Testing for Serial Correlation in Least Squares Regression, I", Biometrika, **37** , 409–
540 428., doi:10.1093/biomet/37.3-4.409. JSTOR 2332391, 1950.

Fathian, F., Mehdizadeh, S., Kozekalani Sales, A., and Safari, M. J. S.: Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. J. Hydrol., 575, 1200–1213, https://doi.org/10.1016/j.jhydrol.2019.06.025, 2019.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., Michaelsen, J.: A quasi-global precipitation time
545 series for drought monitoring: CHIRPS 1981–2012, Clim. Data Rec., 35(23), 121–124, https://doi.org/10.1175/JCLI-D-13-00353.1, 2014.

Ghorbani, M. A., Ahmad Zadeh, H., Isazadeh, M., and Terzi, O.: A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. Environ. Earth Sci., 75, 5096, https://doi.org/10.1007/s12665-015-5096-x, 2016.

550 Jarque, C.M. and Bera, A.K.: Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Economics Letters, 6, 255-259, https://doi.org/10.1016/0165-1765(80)90024-5, 1980.

Katrutsa, A., and Strijov, V.: Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria. Expert Syst. Appl., 76, 1–11, https://doi.org/10.1016/j.eswa.2016.11.005, 2017.

Labat, D., Espinoza, J.C., Ronchail, J., Cochonneau, G., de Oliveira, E., Doudou, J.C., Guyot, J.L.: Fluctuations in the monthly

555   discharge of Guiana Shield rivers related to Pacific and Atlantic climate variability, Hydrol. Sci. J., 57, 1081–1091, https://doi.org/10.1080/02626667.2012.695074, 2012.

Lavery, M. R., Acharya, P., Sivo, S. A., & Xu, L.: Number of predictors and multicollinearity: What are their effects on error and bias in regression? Communications in Statistics - Simulation and Computation, 46(1), 163–183. https://doi.org/10.1080/03610918.2017.1371750, 2017.

560   Xiang, X., Guo, S., Cui, Z., Wang, L., Xu, C. Y.: Improving flood forecast accuracy based on explainable convolutional neural network by Grad-CAM method. Journal of Hydrology, 642, 131867, https://doi.org/10.1016/j.jhydrol.2024.131867, 2024.

Mann, H.B.: Nonparametric tests against trend, Econometrica, 13(3), 245–259, https://doi.org/10.2307/1907187, 1945.

Mesa, O.J., Poveda, G., Hoyos, C.D., Mejía, F., Carvajal, L.F., Cuartas, A., Barco, J.: Predicción no lineal de los caudales medios mensuales del río Nare, Antioquia, Colombia, Av. Recur. Hídricos, 8, 2001.

565   Moeeni, H., Bonakdari, H.: Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model, Stoch. Environ. Res. Risk Assess., 31, 1997–2010, https://doi.org/10.1007/s00477-016-1273-z, 2017.

Nieto, R., Gallego, D., Trigo, R., Ribera, P., Gimeno, L.: Dynamic identification of moisture sources in the Orinoco basin in equatorial South America, Hydrol. Sci. J., 53, 602–617, https://doi.org/10.1623/hysj.53.3.602, 2008.

Niu, W., Feng, Z.: Evaluating the performances of several artificial-intelligence methods in forecasting daily streamflow

570   time-series for sustainable water-resources management, Sustain. Cities Soc., 64, 102564, https://doi.org/10.1016/j.scs.2020.102562, 2021.

Ocampo-Marulanda, C., Fernández-Álvarez, C., Cerón, W. L., Canchala, T., Carvajal-Escobar, Y., Alfonso-Morales, W.: A spatiotemporal assessment of the high-resolution CHIRPS rainfall dataset in southwestern Colombia using combined principal component analysis. Ain Shams Eng. J., 13, 101739, https://doi.org/10.1016/j.asej.2022.101739, 2022.

575   Ocampo-Marulanda, C., Canchala Nastar, T., Carvajal Escobar, Y., Cerón-Arámburo, L.: Estimation of monthly flow data with artificial intelligence in basins with scarce information in the Colombian Orinoquía, Rev. Bras. Recur. Hídricos, 30(2), e14, https://doi.org/10.1590/2318-0331.302520240072, 2025.

Pasquini, A.I., Depetris, P.J.: Discharge trends and flow dynamics of South American rivers draining the southern Atlantic seaboard: An overview, J. Hydrol., 333(2–4), 385–399, https://doi.org/10.1016/j.jhydrol.2006.09.005, 2007.

580   Poveda, G., Jaramillo, A., Gil, M.M., Quiceno, N., Mantilla, R.I.: Seasonality in ENSO-related precipitation, river discharges, soil moisture, and vegetation index in Colombia, Water Resour. Res., 37, 2169–2178, https://doi.org/10.1029/2000WR900395, 2001.

Poveda, G., Mesa, O., Carvajal, L., Hoyos, C., Mejía, J., Cuartas, A., Pulgarin, A.: Predicción de caudales medios mensuales en ríos colombianos usando métodos no lineales, Meteorol. Colomb., 6, 1–8, 2002.

585 Poveda, G., Álvarez, D.M., Rueda, Ó.A.: Hydro-climatic variability over the Andes of Colombia associated with ENSO: A review of climatic processes and their impact on one of the Earth's most important biodiversity hotspots, Clim. Dyn., 36, 2233–2249, https://doi.org/10.1007/s00382-010-0931-y, 2011.

Rafael Miñope, W., Vilcherres Lizárraga, P., Muñoz Pérez, S., Tuesta Monteza, V., Mejía Cabrera, H.: Modelamiento de procesos hidrológicos aplicando técnicas de inteligencia artificial: Una revisión sistemática de la literatura,
590 ITECKNE, 19(1), 46–60, https://doi.org/10.15332/iteckne.v19i1.2645, 2022.

Ruiz-Ochoa, M.A., Vargas-Corredor, Y.A., Orduz-Amaya, L.P., Torres-Corredor, J.S.: Climate variability in water planning in the Cravo Sur River basin (Casanare, Colombia), Inf. Tecnol., 33(4), 117–124, https://doi.org/10.4067/S0718-07642022000400117, 2022.

Scholz, M., Fraunholz, M., and Selbig, J.: Nonlinear principal component analysis: neural network models and applications.
595 En A. N. Gorban, B. Kégl, D. C. Wunsch, and A. Zinovyev (Eds.), Principal manifolds for data visualization and dimension reduction, Springer, Berlin Heidelberg, 58, 44–67, https://doi.org/10.1007/978-3-540-73750-6_2, 2007.

Sen, P.K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63(324), 1379–1389, https://doi.org/10.2307/2285891, 1968.

Silverman, B. W.: Density Estimation for Statistics and Data Analysis. Chapman and Hall/CRC Monographs on Statistics and
600 Applied Probability, 1st edn., Vol. 26, Chapman and Hall, London/New York, 175 pp., https://doi.org/10.1201/9781315140919, 1986

Sirisha, U.M., Belavagi, M.C., Attigeri, G.: Profit prediction using ARIMA, SARIMA and LSTM models in time-series forecasting: A comparison, IEEE Access, 10, 124715–124727, https://doi.org/10.1109/ACCESS.2022.3224938, 2022.

Tootle, G.A., Piechota, T.C., Gutiérrez, F.: The relationships between Pacific and Atlantic Ocean sea-surface temperatures and
605 Colombian streamflow variability, J. Hydrol., 349(3–4), 2008.

Urrea, V., Ochoa, A., Mesa, O.: Validación de la base de datos de precipitación CHIRPS para Colombia a escala diaria, mensual y anual (1981–2014), XXVII Congr. Latinoam. Hidráulica, Lima, Perú, https://www.researchgate.net/publication/310844678, 2016.

Valipour, M.: Long-term runoff study using SARIMA and ARIMA models in the United States, Meteorol. Appl., 22(3), 592–
610 598, https://doi.org/10.1002/met.1491, 2015.

Velásquez, N., Sánchez, J., Carvajal, L.F.: Predicción de caudales medios mensuales en ríos de Colombia usando RNA y MARS, XXIV Congr. Latinoam. Hidráulica, Punta del Este, Uruguay, 2010.

Yao, L., Libera, D.A., Kheimi, M., Sankarasubramanian, A., Wang, D.: The roles of climate forcing and its variability on streamflow at daily, monthly, annual, and long-term scales, Water Resour. Res., 56(7),
615 https://doi.org/10.1029/2020WR027111, 2020.

Yap, Z.N., Musa, S.: Stream-flow forecasting on Pahang River by time-series models ARMA, ARIMA and SARIMA, Recent Trends Civ. Eng. Built Environ., 4(1), 331–341, https://publisher.uthm.edu.my/periodicals/index.php/rtcebe/article/view/5621, 2023.

Wilk, M. B. and Gnanadesikan, R.: Probability plotting methods for the analysis of data. Biometrika, 55, 1–
620   17, https://doi.org/10.1093/biomet/55.1.1, 1968.