

**JEERESD**Journal Home page: www.jeeresd.online

eISSN: 3078-2112



Characterization and Forecasting of Future Drought in the Gambia River Basin at Kédougou Using Standardized Indices, Machine Learning, and Deep Neural Networks

Ibrahima Faye¹, Cheikh Faye^{*1}

1. Department of Geography, Faculty of Science and Technology, Assane Seck University of Ziguinchor, Laboratory of Geomatics and Environment, Ziguinchor, Senegal

ARTICLE INFO

Keywords:

Hydrological drought;
SSI;
SPEI;
SPI;
Climate change;
Gambia River.

ABSTRACT

Background: Drought is one of the most complex and damaging hydroclimatic hazards, particularly in West Africa where water resources are highly sensitive to climate variability and change. This study analyzes the future evolution of meteorological, climatic, and hydrological droughts in the Gambia River basin controlled at the Kédougou gauging station in southeastern Senegal. **Methods:** The approach combines precipitation, potential evapotranspiration (PET), temperature, and simulated streamflow, and computes the SPI, SPEI, and SSI at 1-, 3-, 6-, and 12-month timescales. **Results:**

Results reveal a major contrast between the precipitation signal and thermo-hydrological responses. Under SSP5-8.5, mean temperature increases from 29.67 °C in 2021–2040 to 33.99 °C in 2081–2100, annual PET rises from 1,574 to 1,770 mm, and mean streamflow decreases from 49.02 to 12.45 m³/s. Mean SSI-3 simultaneously deteriorates from –1.21 to –1.94, while the frequency of drought months increases from 67.5% to 91.7% and that of extreme drought months from 35.4% to 84.2%. Under SSP1-2.6, drying remains substantial but is less monotonic. Eight forecasting models are compared at 1-, 3-, and 6-month lead times: Random Forest, ExtraTrees, XGBoost, LightGBM, SVR, MLP/ANN, LSTM, and GRU. LightGBM performs best at one month (NSE = 0.867 under SSP1-2.6 and 0.718 under SSP5-8.5), whereas LSTM/GRU become competitive at longer lead times.

**Corresponding author*

Email: cheikh.faye@univ-zig.sn

Affiliation address: Department of Geography, Faculty of Science and Technology, Assane Seck University of Ziguinchor, Laboratory of Geomatics and Environment, Ziguinchor, Senegal

DOI: 10.68305/jeeresd.v4i4.005

Received: 15 June 2026 **Revised:** 18 July 2026 | **Accepted:** 25 September 2026 | **Published:** 01 October 2026

© 2026 The Authors. Published by EcoClean Environment Company on behalf of JEERESD. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation and reproduction, provided that appropriate credit is given to the original authors and source.



Conclusion : The study highlights a risk of chronic hydrological drought by the end of the century, particularly under high emissions, and demonstrates the value of multimodel forecasting for early warning and adaptive water-resources management.

What this study adds?

1. This study provides an integrated assessment of meteorological, climatic, and hydrological drought propagation using SPI, SPEI, and SSI indices.
2. It reveals future drought intensification under climate change scenarios by quantifying temperature rise, streamflow decline, and increasing drought frequency.
3. This research compares eight artificial intelligence models to improve hydrological drought forecasting at multiple seasonal lead times.
4. It identifies key hydroclimatic predictors controlling forecast performance, supporting early warning systems and adaptive water management strategies.

1. INTRODUCTION

Drought is one of the most complex and damaging hydroclimatic hazards because of its slow onset, potentially long duration, large spatial extent, and capacity to affect simultaneously water resources, agricultural production, ecosystems, energy systems, and socioeconomic activities. Unlike rapidly developing hazards, drought generally results from the progressive accumulation of hydroclimatic anomalies and can propagate through different components of the hydrological cycle (Mishra & Singh, 2010; Wang et al., 2022). A persistent precipitation deficit may first produce meteorological drought before affecting soil moisture, groundwater recharge, river discharge, reservoir storage, and ultimately water availability. Consequently, drought cannot be adequately characterized through precipitation anomalies alone, particularly in regions where rising temperatures and increasing atmospheric evaporative demand may amplify water deficits (Ahmadalipour et al., 2017; Vicente-Serrano et al., 2010). This issue is particularly important in West Africa, where water resources are highly sensitive to climate variability and change. The region experienced severe and persistent drought episodes during the second half of the twentieth century, accompanied by substantial changes in rainfall and river-flow regimes. Although rainfall recovery has been observed in several parts of the Sahel since the late 1990s, strong interannual and spatial variability persists, and hydrological responses do not necessarily follow precipitation trends proportionally (Traore et al., 2014; Rameshwaran et al., 2021). Future climate change could further modify these relationships through changes in precipitation distribution, temperature, evapotranspiration, rainfall extremes, and land-surface processes. Recent climate projections therefore raise important concerns regarding the future reliability of surface-water resources in West African river basins (Bodian et al., 2018; Diop et al., 2025).

The Gambia River basin constitutes a particularly relevant system for investigating these changes. Originating in the Fouta Djallon highlands in Guinea and extending through Senegal and The Gambia, the basin supports agriculture, livestock production, domestic water supply, ecosystems, biodiversity conservation, and numerous socioeconomic activities. Its

hydrological functioning is strongly controlled by the West African monsoon and by the marked seasonal concentration of precipitation (Dione, 1996; Sow, 2007). Historical analyses have demonstrated significant variations and shifts in rainfall and runoff within the basin, highlighting its sensitivity to regional hydroclimatic fluctuations (Traore et al., 2014; Faye & Mendy, 2018). More recent studies have also shown that climate variability, land-use dynamics, and changes in rainfall–runoff relationships may substantially affect the basin's water resources (Kah et al., 2025).

This vulnerability could intensify during the twenty-first century. Climate-model simulations indicate potentially important changes in West African river flows under global warming (Rameshwaran et al., 2021). In the Senegal and Gambia river basins, projected climate changes have already been associated with substantial modifications of future streamflow (Bodian et al., 2018). More specifically, recent assessments of the Gambia River basin indicate that changes in precipitation, temperature, and evapotranspiration may translate into significant changes in water availability and hydrological regimes (Séne et al., 2024). Such projections emphasize the need to assess drought not merely as a meteorological phenomenon but as a process propagating from climatic forcing to hydrological response.

Several standardized indices have been developed to characterize these complementary dimensions of drought. The Standardized Precipitation Index (SPI) is widely used to quantify meteorological drought through precipitation anomalies accumulated over different timescales. However, precipitation-based indices may underestimate drought intensification under warming conditions because they do not explicitly account for increasing atmospheric evaporative demand. The Standardized Precipitation Evapotranspiration Index (SPEI) addresses this limitation by incorporating the climatic water balance between precipitation and potential evapotranspiration and is therefore particularly suited to drought assessment in a warming climate (Vicente-Serrano et al., 2010). At the hydrological level, standardized runoff or streamflow indices provide direct information on anomalies in river discharge and water availability (Shukla & Wood, 2008). The Standardized Streamflow Index (SSI), in particular, enables hydrological drought to be characterized consistently across different temporal scales (Vicente-Serrano et al., 2012). Joint analysis of SPI, SPEI, and SSI can therefore provide a more comprehensive understanding of drought propagation from atmospheric deficits to river-flow responses.

The importance of such an integrated framework increases under future climate conditions because drought propagation is nonlinear and influenced by multiple interacting controls. Temperature rise can intensify evapotranspiration and modify the conversion of rainfall into runoff, meaning that relatively moderate precipitation changes may generate disproportionately large hydrological responses (Ahmadalipour et al., 2017; Wang et al., 2022). Land-use and land-cover changes can further alter infiltration, runoff generation, and basin storage, thereby modifying hydrological drought characteristics (Kah et al., 2025; Liu et al., 2025). Moreover, climate projections themselves contain uncertainties related to emission pathways, climate-model structure, and the representation of regional hydroclimatic processes (Abdulai & Chung, 2019; Song et al., 2021). Assessing drought under contrasting Shared

Socioeconomic Pathways is consequently important for distinguishing the implications of lower- and higher-emission futures. The SSP framework provides internally consistent trajectories of socioeconomic development and greenhouse-gas forcing (Riahi et al., 2017), while ScenarioMIP provides the coordinated climate-model framework used to investigate these alternative futures (O'Neill et al., 2016).

Beyond drought characterization, reliable forecasting is essential for translating climate information into operational early-warning and water-management decisions. Traditional statistical and process-based hydrological approaches remain fundamental, but the increasing availability of hydroclimatic datasets has stimulated the development of machine-learning and deep-learning methods capable of representing complex nonlinear relationships. Random Forest, for example, provides a flexible ensemble-learning framework capable of handling nonlinear interactions among numerous predictors (Breiman, 2001). Recurrent neural-network architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models are particularly promising because they can learn temporal dependencies and hydrological memory from sequential data. Recent reviews demonstrate the rapidly increasing application of deep-learning techniques to streamflow, hydrological, and drought forecasting (Zhao et al., 2024). Nevertheless, predictive performance depends strongly on forecast horizon, input-variable selection, hydroclimatic regime, and the stationarity of relationships between predictors and drought indicators.

Despite these methodological advances, relatively few studies have jointly investigated future meteorological, climatic, and hydrological drought while simultaneously evaluating the predictability of hydrological drought at several lead times in the upper Gambia River basin. Existing research has substantially improved understanding of historical hydroclimatic variability, rainfall–runoff relationships, future streamflow, and climate-change impacts in the basin (Faty et al., 2018; Faty et al., 2023; Séne et al., 2024; Kah et al., 2025), but an integrated assessment combining multiple standardized drought indices with machine-learning and deep-learning forecasting remains necessary. Such an approach is especially relevant at Kédougou, where the upstream catchment provides an important hydrological control point for understanding climate-induced changes in the continental Gambia River system.

Accordingly, this study aims to characterize and forecast future drought in the Gambia River basin at the Kédougou gauging station under contrasting climate-change scenarios. Specifically, it seeks to: (i) quantify future meteorological, climatic, and hydrological drought using SPI, SPEI, and SSI at multiple timescales; (ii) examine the evolution of drought intensity, frequency, duration, and severity under SSP1-2.6 and SSP5-8.5 relative to the 1985–2014 reference climatology; (iii) evaluate eight machine-learning and deep-learning models—Random Forest, ExtraTrees, XGBoost, LightGBM, SVR, MLP/ANN, LSTM, and GRU—for forecasting SSI-3 at 1-, 3-, and 6-month lead times; and (iv) identify the principal hydroclimatic predictors controlling forecast skill. By integrating climate-change projections, standardized drought indicators, and artificial-intelligence-based forecasting, the study seeks to provide a scientifically robust framework for drought early warning and adaptive water-resources management in the Gambia River basin.

Novelty and Contribution Relative to Previous Work :

Several studies have substantially improved understanding of hydroclimatic variability, rainfall–runoff relationships, and climate-change impacts in the Gambia River basin. Historical analyses have documented significant variations in rainfall and runoff (Traore et al., 2014; Faye & Mendy, 2018), while recent assessments have projected future streamflow changes under climate forcing (Bodian et al., 2018; Séne et al., 2024; Kah et al., 2025). However, these previous studies have generally focused on either historical variability or future streamflow projections in isolation, without jointly examining meteorological, climatic, and hydrological drought propagation using multiple standardized indices.

This study provides four principal novelties that distinguish it from previous work on the Gambia River basin:

First, we simultaneously characterize meteorological drought (SPI), climatic drought (SPEI), and hydrological drought (SSI) using a consistent framework and a fixed 1985–2014 reference climatology. This integrated approach reveals the propagation and amplification of drought signals through the hydrological system—a dimension that has not been systematically analyzed in the Gambia basin.

Second, we evaluate future drought dynamics under two contrasting Shared Socioeconomic Pathways (SSP1-2.6 and SSP5-8.5), whereas previous projections have typically used older Representative Concentration Pathways (RCPs) or single-scenario analyses. The scenario contrast allows us to quantify the benefits of emission mitigation for water-resource sustainability.

Third, we compare eight machine-learning and deep-learning models (Random Forest, ExtraTrees, XGBoost, LightGBM, SVR, MLP/ANN, LSTM, and GRU) for forecasting SSI-3 at multiple lead times (1, 3, and 6 months). No previous study has systematically evaluated such a comprehensive suite of artificial-intelligence models for drought forecasting in the Gambia basin.

Fourth, we identify the principal hydroclimatic predictors controlling forecast skill, revealing a transition from hydrological persistence at short lead times to multi-timescale hydroclimatic control at seasonal horizons. This finding has direct operational implications for early-warning system design.

By integrating climate projections, standardized drought indicators, and artificial-intelligence-based forecasting, this study provides a scientifically robust framework that extends beyond previous work and offers practical guidance for drought early warning and adaptive water-resources management in the Gambia River basin.

2. MATERIALS AND METHODS**2.1 Study Area**

The Gambia River basin is one of the major river systems in West Africa. It covers approximately 77,100 km² across three countries: Guinea (15.25% of the basin area), Senegal (72.30%), and The Gambia (12.45%). The Gambia River rises in the Fouta Djallon highlands in Guinea, at an elevation of approximately 1,150 m near Labé, before flowing nearly 1,180 km to

its mouth in the Atlantic Ocean in The Gambia (Kah et al., 2025). It comprises two main sections: a continental reach and a maritime reach (Dione, 1996; Sow, 2007; Faye & Mendy, 2018). The continental reach receives numerous tributaries on its left bank, including the Diagueri, Niokolo-Koba, Niéri-Ko, and Sandougou, and on its right bank, notably the Thiokoye, Diarha, and Koulountou (Figure 1).

Its hydrological regime is tropical Sudano-Guinean and is strongly influenced by the West African monsoon, with a marked alternation between a long dry season (November to May) and a rainy season (June to October). Streamflow is mainly rainfall-driven, with precipitation concentrated between July and September, when the principal floods occur (Séne et al., 2024). The Kédougou gauging station, located in southeastern Senegal, controls a catchment of approximately 7,550 km² and therefore represents an important part of the upper Gambia River basin. The area is characterized by relatively rugged topography dominated by the foothills of the Fouta Djallon, Precambrian basement geological formations, and vegetation consisting mainly of wooded savannas and open forests. Annual rainfall, generally ranging from 1,000 to 1,400 mm, is among the highest in Senegal and explains the substantial streamflow recorded at the station. Discharge shows strong seasonal variability, with pronounced low flows during the dry season and rapid floods during the wet season.

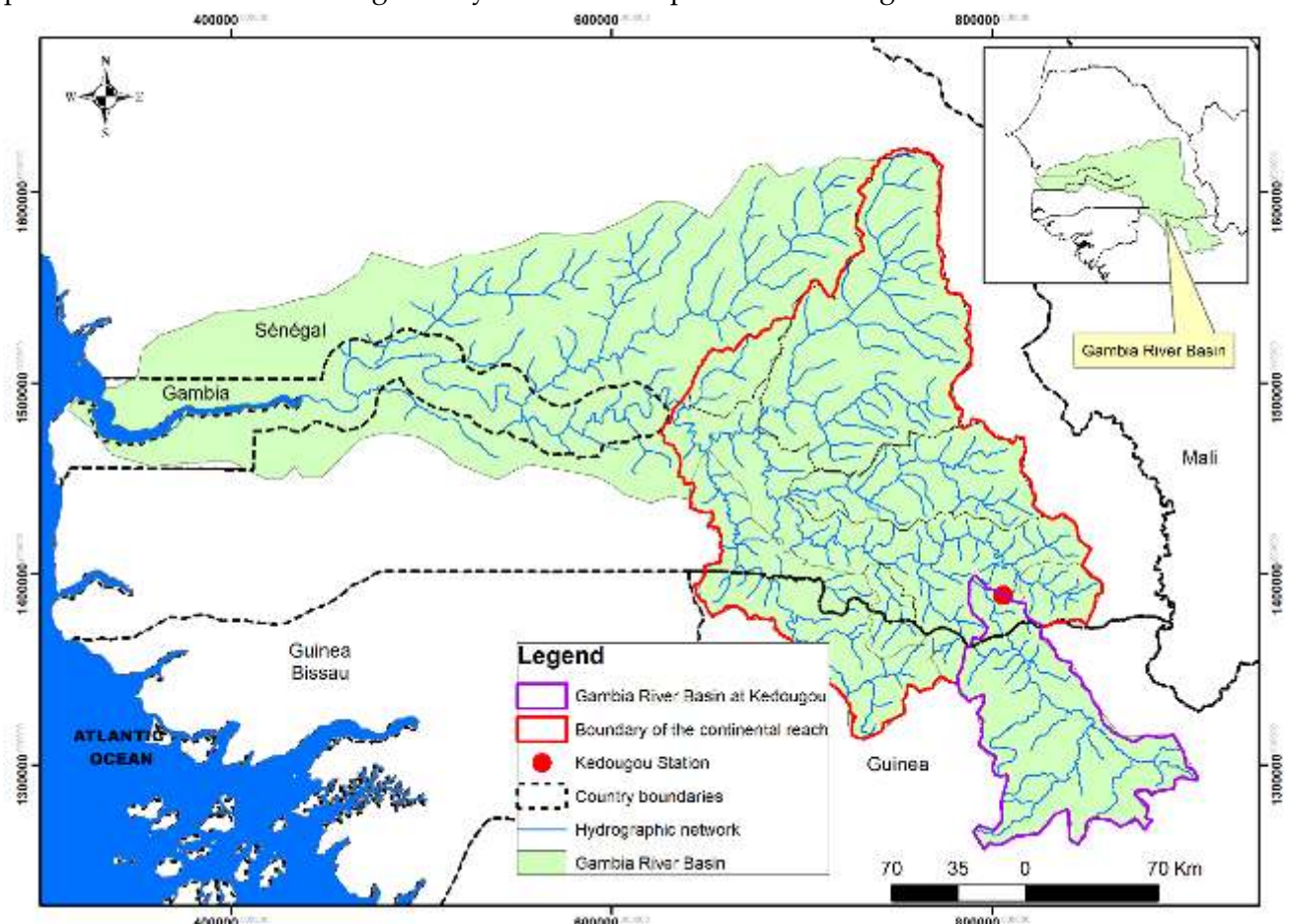


Figure 1: Gambia River basin at the Kédougou station

Because of its upstream position and the quality of its hydrometric records, the Kédougou station is a reference site for investigating hydrological variability, climate-change impacts on water resources, rainfall–runoff modeling, and river-flood risk in the Gambia River basin.

2.2 Hydroclimatic data, climate forcing and climate-model processing

Future climate forcing was obtained from the ISIMIP3b climate-forcing archive. The analysis considered the CMIP6 general circulation models MPI-ESM1-2-HR and MRI-ESM2-0 under two contrasting emission pathways, SSP1-2.6 and SSP5-8.5. For both models, the r1i1p1f1 ensemble-member realization was used. Daily precipitation and air-temperature variables were extracted at the $0.5^\circ \times 0.5^\circ$ spatial resolution distributed by ISIMIP3b. The ISIMIP forcing used in this study had already undergone bias adjustment and statistical downscaling through the ISIMIP3BASD procedure, using W5E5 as the observational reference dataset. No additional bias correction or statistical downscaling was applied after acquisition of the ISIMIP3b data.

For each GCM–scenario combination, grid cells intersecting the Kédougou upstream catchment were extracted and averaged over the basin. Daily precipitation and PET were subsequently aggregated to monthly totals, whereas temperature and discharge were aggregated as monthly means. The two GCM simulations were then combined using an arithmetic ensemble mean. All drought indices were computed separately from the resulting scenario-specific series using 1985–2014 as the fixed historical reference climatology.

The use of two GCMs was intended to represent part of the climate-model structural uncertainty. Nevertheless, this limited ensemble does not encompass the full range of CMIP6 uncertainty and projections should therefore be interpreted as conditional on the selected climate models and emission scenarios.

2.2.1 Historical rainfall–runoff model and future streamflow simulation

Model hyperparameters were optimized using RandomizedSearchCV combined with time-series cross-validation (TimeSeriesSplit), thereby preserving the chronological ordering of observations during model selection. The Random Forest estimator was configured with bootstrap sampling enabled (bootstrap=True), parallel processing using all available CPU cores (n_jobs=-1), and a fixed random seed (random_state=42) to ensure reproducibility. Hyperparameter optimization consisted of 40 randomly sampled configurations (n_iter=40) evaluated using negative mean squared error as the optimization criterion.

The hyperparameter search space comprised the number of trees (n_estimators = 400, 600, 800, 1000, or 1200), maximum tree depth (max_depth = 8, 10, 12, 15, 18, or unlimited), minimum number of samples required to split an internal node (min_samples_split = 4, 6, 8, 10, 15, or 20), minimum number of samples required at a terminal leaf (min_samples_leaf = 3, 5, 7, 10, 15, or 20), number or proportion of predictors considered at each split (max_features = 0.5, 0.7, 0.9, 1.0, or "sqrt"), and proportion of the training sample used to grow each tree (max_samples = 0.6, 0.7, 0.8, 0.9, or 1.0). Five temporal cross-validation splits were specified whenever permitted by the calibration-sample size. The best-performing hyperparameter combination

according to the cross-validated negative mean squared error was automatically retained and refitted to the complete calibration dataset.

The resulting optimal Random Forest configuration was `n_estimators = 1000`, `max_depth = 18`, `min_samples_split = 6`, `min_samples_leaf = 20`, `max_features = "sqrt"`, `max_samples = 0.8`, `bootstrap = True`, and `random_state = 42`.

Before model fitting, observed discharge values were constrained to non-negative values and transformed using the natural logarithmic transformation $\log(1 + Q)$ ($\log1p$). Model predictions were subsequently back-transformed to the original discharge scale using `expm1` and constrained to non-negative values. This transformation was introduced to reduce the disproportionate influence of extreme flood values during model fitting while maintaining physically meaningful non-negative discharge predictions.

The observed hydroclimatic dataset was sorted chronologically before model development, and duplicate dates were removed. Climatic and hydroclimatic-memory predictors were subsequently generated, after which observations containing missing values in either the predictor set or observed discharge were excluded. The final model contained 89 climatic and hydroclimatic-memory predictor variables. Importantly, no lagged discharge variable (`Debit_lag_n`) was included in the predictor set. Future discharge projections were therefore driven exclusively by climatic forcing and climate-derived antecedent-memory indicators and did not recursively depend on previously simulated discharge values.

The calibration–validation separation was strictly chronological rather than random. After preprocessing and construction of the complete predictor matrix, the executed modelling workflow retained 3,285 daily observations for calibration, covering 1 January 1992 to 31 December 2000, and 2,190 daily observations for independent validation, covering 1 January 2001 to 31 December 2006. The validation observations therefore occurred temporally after the calibration observations, and no random shuffling was performed. This temporal separation, together with `TimeSeriesSplit` during hyperparameter optimization, was designed to minimize temporal information leakage and provide an independent assessment of model generalization to later hydroclimatic conditions.

Model performance was evaluated separately for the calibration and independent validation subsets using the Nash–Sutcliffe efficiency (NSE), Kling–Gupta efficiency (KGE), root-mean-square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), percent bias (PBIAS), and Pearson correlation coefficient. During calibration, the model achieved $\text{NSE} = 0.878$, $\text{KGE} = 0.808$, $R^2 = 0.878$, $\text{RMSE} = 50.56 \text{ m}^3 \text{ s}^{-1}$, $\text{MAE} = 19.10 \text{ m}^3 \text{ s}^{-1}$, $\text{PBIAS} = -8.33\%$, and $r = 0.944$. On the independent validation dataset, performance remained satisfactory, with $\text{NSE} = 0.649$, $\text{KGE} = 0.645$, $R^2 = 0.649$, $\text{RMSE} = 94.68 \text{ m}^3 \text{ s}^{-1}$, $\text{MAE} = 48.87 \text{ m}^3 \text{ s}^{-1}$, $\text{PBIAS} = -23.70\%$, and $r = 0.835$.

The reduction in performance between calibration and validation indicates some loss of predictive skill when the model is transferred to an independent later period, particularly in terms of discharge magnitude, as reflected by the negative validation PBIAS. Nevertheless, the positive NSE and KGE values and the strong validation correlation indicate that the model retained meaningful out-of-sample predictive capability. Overfitting was further controlled

through chronological data separation, time-series cross-validation, randomized hyperparameter optimization, restrictions on tree depth and terminal-leaf size, and independent validation on observations not used for model fitting or hyperparameter selection.

The representativeness of the two subsets was evaluated by comparing annual precipitation totals, mean discharge conditions, seasonal variability, and the occurrence of relatively wet and dry years. Wet and dry years were classified according to annual precipitation relative to the long-term mean annual precipitation calculated over the complete 1992–2006 reference period. The mean annual precipitation over this period was approximately 1,199.65 mm yr⁻¹. Years with annual precipitation greater than or equal to this value were classified as wet years, whereas years with precipitation below this threshold were classified as dry years.

According to this criterion, the 1992–2000 calibration period contained four wet years (1994, 1997, 1999, and 2000) and five dry years (1992, 1993, 1995, 1996, and 1998). The 2001–2006 validation period contained three wet years (2001, 2002, and 2003) and three dry years (2004, 2005, and 2006). Both calibration and validation subsets therefore include contrasting hydroclimatic conditions rather than representing exclusively wet or dry phases. This distribution supports the use of the validation period as an independent test of the model's ability to reproduce discharge variability under conditions differing from those encountered during calibration.

2.2.2 Transfer of the rainfall–runoff model to future climate conditions

The calibrated RF model was transferred to future climate conditions under the assumption that the empirical relationships learned between hydroclimatic forcing and discharge remain sufficiently informative under the projected climate range. This assumption does not imply strict hydrological stationarity. Rather, the model is used as a statistical rainfall–runoff transfer function conditional on precipitation, temperature, PET, antecedent conditions and seasonal predictors.

Future simulations were initialized from 2015, with 2015–2020 treated as a warm-up period to construct lagged and rolling predictors before analysis began in 2021. Thereafter, daily discharge was generated sequentially until 2100 for each climate-model–scenario combination. An important limitation is that Random Forest models are interpolation-oriented algorithms and do not extrapolate physical relationships beyond the range represented in their training data in the same manner as process-based hydrological models. Consequently, projected discharge becomes increasingly uncertain when future combinations of precipitation, PET and temperature depart substantially from the historical predictor space. We therefore interpret late-century discharge and SSI estimates as scenario-conditioned projections rather than deterministic forecasts. This extrapolation uncertainty is expected to be greatest under SSP5-8.5.

2.3 Calculation of standardized drought indices

SPI, SPEI and SSI were calculated at 1-, 3-, 6-, and 12-month accumulation timescales. A fixed historical climatological period extending from January 1985 to December 2014 (360 months) was used exclusively for calibration and standardization of the drought indices. Data from 2015 onward were excluded from the reference distribution. Monthly precipitation and PET were obtained as monthly totals, whereas temperature and discharge were represented by monthly means.

Standardization was performed separately for each calendar month and accumulation timescale in order to preserve the marked seasonality of the basin. For SPI, precipitation was accumulated over 1-, 3-, 6-, and 12-month moving windows. For SPEI, the monthly climatic water balance, defined as precipitation minus potential evapotranspiration ($P - PET$), was accumulated over the same timescales. For SSI, monthly mean discharge was averaged over moving windows of 1, 3, 6, and 12 months.

For each calendar month and accumulation timescale, the corresponding values from the 1985-2014 reference period were ranked and converted to empirical non-exceedance probabilities using the Gringorten plotting-position formula:

$$p_i = (i - 0.44) / (n + 0.12)$$

where i denotes the ordered rank and n is the number of available observations in the corresponding calendar-month reference sample. Tied observations were assigned their average rank. The resulting empirical non-exceedance probabilities were subsequently transformed into standardized normal scores using the inverse cumulative distribution function of the standard normal distribution, $Z = \Phi^{-1}(p)$. This empirical normal-score procedure was applied consistently to precipitation, climatic water balance and streamflow and therefore did not require fitting a predefined parametric probability distribution.

For SPI, zero precipitation was explicitly accounted for because rain-free conditions may occur during the dry season. The probability of zero precipitation, q_0 , was estimated as the proportion of zero values in the 1985-2014 reference sample for the corresponding calendar month and accumulation timescale. For zero precipitation, the probability was assigned to the midpoint of the zero-probability mass ($q_0/2$). For strictly positive precipitation values, the empirical probability $G(x)$, obtained from the positive reference observations, was adjusted according to $H(x) = q_0 + (1 - q_0)G(x)$. This treatment retained zero-precipitation months as valid observations and prevented their exclusion from the empirical reference distribution.

For future conditions, the 1985-2014 empirical distributions were kept fixed throughout the 2021-2100 projection period. Future observations were therefore standardized relative to the same historical climatology rather than being recalibrated separately for each future period or emission scenario. This ensures that changes in SPI, SPEI, and SSI represent departures from a common historical reference and allows direct comparison between SSP1-2.6 and SSP5-8.5.

For projected values located between historical observations, empirical probabilities were obtained by linear interpolation between adjacent reference plotting positions. Values below the minimum of the 1985-2014 reference distribution were assigned the empirical probability associated with the historical minimum, whereas values exceeding the historical maximum

were assigned the probability associated with the historical maximum. Thus, no parametric tail extrapolation was applied. Before transformation with the inverse standard-normal function, probabilities were numerically constrained to the interval $[10^{-6}, 1 - 10^{-6}]$ to prevent undefined infinite standardized scores. In practice, for values outside the historical range, the effective bounds were generally determined by the endpoint Gringorten plotting positions of the corresponding monthly empirical distribution.

No future observations were incorporated into the estimation of the historical reference distribution. Consequently, projected drought-index values quantify the relative rarity of future hydroclimatic conditions with respect to the fixed 1985–2014 baseline.

Drought conditions were primarily identified using SSI-3, with drought defined as $\text{SSI-3} \leq -1.0$, severe drought as $\text{SSI-3} \leq -1.5$, and extreme drought as $\text{SSI-3} \leq -2.0$. A drought event was defined as a continuous sequence of months satisfying the $\text{SSI-3} \leq -1.0$ threshold. SSI-3 was subsequently retained as the principal predictand for the artificial-intelligence forecasting experiments.

2.4 Artificial-intelligence forecasting and reproducibility

Eight machine-learning and deep-learning algorithms were evaluated for forecasting the 3-month Standardized Streamflow Index (SSI-3): Random Forest (RF), Extra Trees (ET), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Support Vector Regression (SVR), Multilayer Perceptron/Artificial Neural Network (MLP/ANN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU).

The predictor set comprised monthly precipitation, potential evapotranspiration (PET), temperature, discharge, SPI-3, SPEI-3, SSI-1, SSI-3, and SSI-6. For the conventional machine-learning algorithms, predictor values were represented at lags of 0, 1, 2, 3, 6, and 12 months. Here, lag 0 represents information available at the forecast-origin month and does not correspond to information from the forecast target month. For the recurrent deep-learning models (LSTM and GRU), predictors were arranged as sequences covering the preceding 12 months. Forecasts were generated for lead times of 1, 3, and 6 months.

All preprocessing and modelling were implemented in Python. The workflow relied on NumPy and pandas for numerical and tabular data processing, scikit-learn for conventional machine-learning algorithms and performance evaluation, XGBoost and LightGBM for gradient-boosting models, and a deep-learning framework for implementation of the LSTM and GRU architectures. Exact Python, package, and deep-learning-framework versions should be reported from the computational environment used for the final model runs to ensure full software reproducibility.

Prior to modelling, dates and numerical fields were checked for consistency and daily hydroclimatic data were aggregated to a monthly time step. Precipitation and PET were aggregated as monthly totals, whereas temperature and discharge were represented by monthly means. SPI, SPEI, and SSI were then calculated at 1-, 3-, 6-, and 12-month accumulation scales using the fixed 1985–2014 historical climatology. Rows for which the

required lagged predictors or forecast targets could not be constructed were excluded from model fitting rather than filled using future information.

Predictor preprocessing was performed within the temporal training framework to prevent information leakage. Tree-based algorithms were fitted without requiring feature scaling, whereas algorithms sensitive to predictor magnitude, including SVR, MLP/ANN, LSTM, and GRU, were standardized using scaling parameters estimated from the training data. The same training-derived transformations were subsequently applied to validation and test observations, without refitting preprocessing parameters on either dataset.

Model development followed a strictly chronological train-validation-test design and was performed separately for SSP1-2.6 and SSP5-8.5. Forecast targets up to 2050 were used for model training, the 2051–2075 period was reserved for validation and model selection, and 2076–2100 constituted an independent test period. No random shuffling across these temporal subsets was performed. This design ensured that model evaluation reproduced the operational forecasting situation in which models trained on earlier observations are applied to later, unseen hydroclimatic conditions.

Hyperparameter optimization and model selection were performed exclusively using the training and validation periods. The independent 2076–2100 test period was not used for feature selection, hyperparameter optimization, preprocessing-parameter estimation, or model selection. Final model performance was therefore evaluated on observations that were temporally and computationally independent of model development.

For the neural-network models, LSTM and GRU architectures were trained using 12-month input sequences. The exact number of hidden units, number of layers, dropout rate, optimizer, learning rate, batch size, maximum number of epochs, early-stopping patience, and random seed should be reported from the final training script or saved model configuration. These parameters cannot be reconstructed reliably from prediction outputs alone and should not be inferred retrospectively.

Continuous SSI-3 forecasts were evaluated using root-mean-square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), Pearson correlation coefficient (r), Nash–Sutcliffe efficiency (NSE), and Kling–Gupta efficiency (KGE). Drought classification performance, using $SSI-3 \leq -1.0$ as the drought threshold, was assessed using accuracy, precision, recall, and F1 score. Model performance was evaluated separately for each emission scenario and forecast horizon.

The final independent test results demonstrate that the identity of the best-performing model depended on both scenario and forecast lead time. Under SSP1-2.6, LightGBM provided the best performance at the 1-month horizon (NSE = 0.867, KGE = 0.870, RMSE = 0.313, F1 = 0.960), LSTM at 3 months (NSE = 0.713, KGE = 0.671, RMSE = 0.459, F1 = 0.951), and GRU at 6 months (NSE = 0.539, KGE = 0.591, RMSE = 0.582, F1 = 0.887). Under SSP5-8.5, LightGBM performed best at 1 month (NSE = 0.718, KGE = 0.693, RMSE = 0.308, F1 = 0.978), whereas GRU and LSTM provided the strongest results at the 3- and 6-month horizons, respectively. These results indicate that no single algorithm was universally optimal across scenarios and forecast lead times.

2.4.1 Predictor–target alignment and prevention of information leakage

Special attention was given to predictor–target alignment to prevent temporal information leakage. For each forecast issued at month t and lead time h , where $h = 1, 3$, or 6 months, the predictor vector contained only information available at or before the forecast-origin month t . A predictor labelled “lag 0” therefore referred to the contemporaneous value available at month t and not to a value observed at the future target month $t+h$.

Formally, the forecasting problem was defined as:

$$X(t) = \{x_j(t), x_j(t-1), x_j(t-2), x_j(t-3), x_j(t-6), x_j(t-12)\},$$

where x_j represents the hydroclimatic predictors, and

$$y(t,h) = \text{SSI-3}(t+h), \quad h \in \{1, 3, 6\} \text{ months..}$$

Thus, for a one-month-ahead forecast, predictors available through month (t) were used to predict SSI-3 at $(t+1)$; for the three-month horizon, the same information set was used to predict SSI-3 at $(t+3)$; and for the six-month horizon, SSI-3 at $(t+6)$ was predicted. No predictor dated later than the forecast origin (t) was included in $(X(t))$.

The predictor set comprised precipitation, potential evapotranspiration (PET), temperature, discharge, SPI-3, SPEI-3, SSI-1, SSI-3, and SSI-6. For the conventional machine-learning algorithms, predictors were represented at lags of 0, 1, 2, 3, 6, and 12 months. For LSTM and GRU, the input consisted of a 12-month sequence ending at the forecast-origin month (t) . Consequently, the most recent element of each recurrent-model input sequence corresponded to month (t) , whereas the target remained SSI-3 at $(t+h)$.

All accumulated and standardized drought indices used as predictors were constructed using information available up to and including their corresponding predictor month. For example, SSI-3 at month (t) was calculated from discharge information ending at month (t) , and no discharge observation from $(t+1)$ or later contributed to that predictor. The same temporal constraint was applied to SPI-3, SPEI-3, SSI-1, and SSI-6.

For each forecast horizon, the target series was generated by shifting SSI-3 forward by (h) months relative to the predictor matrix. Predictor–target pairs for which either the lagged predictors or future target could not be constructed were removed before model fitting. This alignment ensured that each row of the modelling dataset represented information available at a specific forecast origin and a target occurring strictly in the future.

The modelling dataset was then divided chronologically and independently for SSP1-2.6 and SSP5-8.5. Forecast targets up to December 2050 were assigned to the training period, targets from January 2051 to December 2075 to the validation period, and targets from January 2076 to December 2100 to the independent test period. Samples were assigned to these subsets according to the date of the forecast target rather than through random sampling.

To further prevent boundary leakage, predictor–target pairs were retained only when the shifted target belonged entirely to the intended temporal subset. Thus, a predictor observation located near the end of the training period whose $(t+h)$ target fell within the validation period was not treated as a training target. The same principle was applied at the validation–test boundary.

No random shuffling was performed across the training, validation, and test periods. Hyperparameter selection and model comparison were conducted using only the training and validation data. The 2076–2100 test period remained isolated from model development and was used only for final out-of-sample evaluation. Similarly, preprocessing parameters requiring estimation from data, including feature-scaling parameters for scale-sensitive algorithms, were estimated from the training data and subsequently applied unchanged to validation and test data.

This forecasting design therefore prevented three major forms of information leakage: (i) direct use of predictor values occurring after the forecast origin, (ii) contamination caused by shifted targets crossing temporal dataset boundaries, and (iii) estimation of preprocessing or model-selection parameters using the independent test period.

2.4.2 Baseline models and forecast evaluation

To determine whether the eight machine-learning and deep-learning models provided predictive skill beyond simple forecasting rules, their performance was evaluated relative to persistence, climatological, and naïve classification benchmarks.

The persistence benchmark assumed that the SSI-3 value available at the forecast origin remained unchanged over the forecast horizon:

$$\text{SSI-3}_{\text{hat}}(t+h, \text{persistence}) = \text{SSI-3}(t), \quad h \in \{1, 3, 6\}.$$

Thus, the same SSI-3 value observed or simulated at month t was used as the prediction for $t+1$, $t+3$, or $t+6$. Persistence provides a particularly relevant benchmark for drought forecasting because standardized hydrological indices exhibit temporal autocorrelation, especially at multi-month accumulation scales. A complex forecasting model should therefore outperform persistence to demonstrate predictive information beyond the inherent memory of SSI-3.

A climatological benchmark was also defined using the fixed 1985–2014 historical reference climatology. For each calendar month m , the benchmark prediction was the mean SSI-3 value calculated from the corresponding calendar month during the 1985–2014 reference period:

$$\text{SSI-3}_{\text{hat_clim}}(m) = (1/N_m) \sum \text{SSI-3}(k,m),$$

where N_m is the number of valid SSI-3 observations available for calendar month m during 1985–2014. The climatological forecast for a target month therefore depended only on its calendar month and not on future scenario information. Because SSI-3 was standardized separately by calendar month relative to the 1985–2014 reference period, this benchmark is expected to remain close to zero, but it was calculated explicitly from the historical reference values rather than assumed to be exactly zero.

Continuous forecast performance was quantified using root-mean-square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), Pearson correlation coefficient (r), Nash–Sutcliffe efficiency (NSE), and Kling–Gupta efficiency (KGE). These complementary statistics quantify different aspects of predictive performance, including absolute error magnitude, linear association, explained variance, overall predictive efficiency, bias, and representation of variability.

The machine-learning and deep-learning models were compared directly with the persistence and climatological benchmarks at each forecast horizon and under each emission scenario. This comparison was used to determine whether increased model complexity resulted in genuine out-of-sample predictive improvement rather than merely reproducing climatological conditions or the temporal persistence of SSI-3.

For categorical evaluation, drought occurrence was defined as $\text{SSI-3} \leq -1.0$. Each continuous SSI-3 prediction was therefore converted into a binary drought/non-drought classification using the same threshold. Classification performance was initially assessed using accuracy, precision, recall (sensitivity), and F1 score.

Because drought prevalence becomes highly imbalanced during parts of the future simulations, particularly under SSP5-8.5, accuracy and F1 score alone were not considered sufficient to characterize classification performance. Additional metrics were therefore defined to account explicitly for both drought and non-drought classes.

Balanced accuracy was calculated as the arithmetic mean of sensitivity and specificity:

$$\text{Balanced Accuracy} = (\text{Sensitivity} + \text{Specificity}) / 2,$$

where Sensitivity = $\text{TP}/(\text{TP}+\text{FN})$ and Specificity = $\text{TN}/(\text{TN}+\text{FP})$. Here, TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

The Matthews correlation coefficient (MCC) was calculated as:

$$\text{MCC} = (\text{TP} \times \text{TN} - \text{FP} \times \text{FN}) / \sqrt{[(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})]}.$$

MCC ranges from -1 to $+1$ and remains informative even when the two classes occur at substantially different frequencies. A value of $+1$ indicates perfect classification, 0 indicates no overall correlation between predictions and observations, and -1 indicates complete disagreement.

Cohen's kappa was additionally used to quantify classification agreement after accounting for agreement expected by chance. Precision–recall area under the curve (PR-AUC) was used to characterize the trade-off between precision and recall across classification thresholds and was considered particularly informative for imbalanced drought occurrence. Receiver operating characteristic area under the curve (ROC-AUC) was additionally calculated when continuous prediction scores permitted threshold-independent evaluation.

Confusion matrices were generated for each model, scenario, and forecast horizon to explicitly report TP, TN, FP, and FN counts and to facilitate interpretation of the classification metrics. Finally, classification performance was compared with a naïve majority-class benchmark. For each modelling experiment, the majority class was determined exclusively from the corresponding training dataset. The naïve classifier subsequently assigned that class to every validation or test observation without using any predictor information. This benchmark provides the minimum reference needed to determine whether apparently high accuracy under strongly imbalanced drought conditions reflects genuine discriminatory skill or simply dominance of the most frequent class.

All baseline forecasts and evaluation metrics were computed separately for SSP1-2.6 and SSP5-8.5 and independently for the 1-, 3-, and 6-month forecast horizons. The final independent

comparison was performed on the 2076–2100 test period, which was not used during model fitting, hyperparameter tuning, feature selection, or preprocessing-parameter estimation.

3. RESULTS

3.1 Future Hydroclimatic Changes

Under SSP1-2.6, annual precipitation does not follow a monotonic trend: it reaches 1,015.5 mm yr⁻¹ in 2021–2040 and 1,111.0 mm yr⁻¹ in 2041–2060, decreases to 925.4 mm yr⁻¹ in 2061–2080, and rises again to 1,101.5 mm yr⁻¹ in 2081–2100 (Table 1). In contrast, PET increases from 1,574.4 to 1,604.0 mm yr⁻¹ and temperature from 29.67 to 30.32 °C. Mean streamflow decreases from 44.41 to 35.21 m³/s between the beginning and end of the century, with a minimum of 33.59 m³/s in 2061–2080.

Table 1. Future hydroclimatic statistics by scenario and period.

Scenario	Period	P (mm/yr)	PET (mm/yr)	T (°C)	Mean Q (m ³ /s)
SSP1-2.6	2021-2040	1015.5	1574.4	29.67	44.41
SSP1-2.6	2041-2060	1111.0	1588.0	29.97	46.49
SSP1-2.6	2061-2080	925.4	1602.9	30.30	33.59
SSP1-2.6	2081-2100	1101.5	1604.0	30.32	35.21
SSP5-8.5	2021-2040	1101.6	1574.5	29.67	49.02
SSP5-8.5	2041-2060	943.1	1634.3	30.99	27.57
SSP5-8.5	2061-2080	1063.9	1703.4	32.51	17.92
SSP5-8.5	2081-2100	994.6	1770.3	33.99	12.45

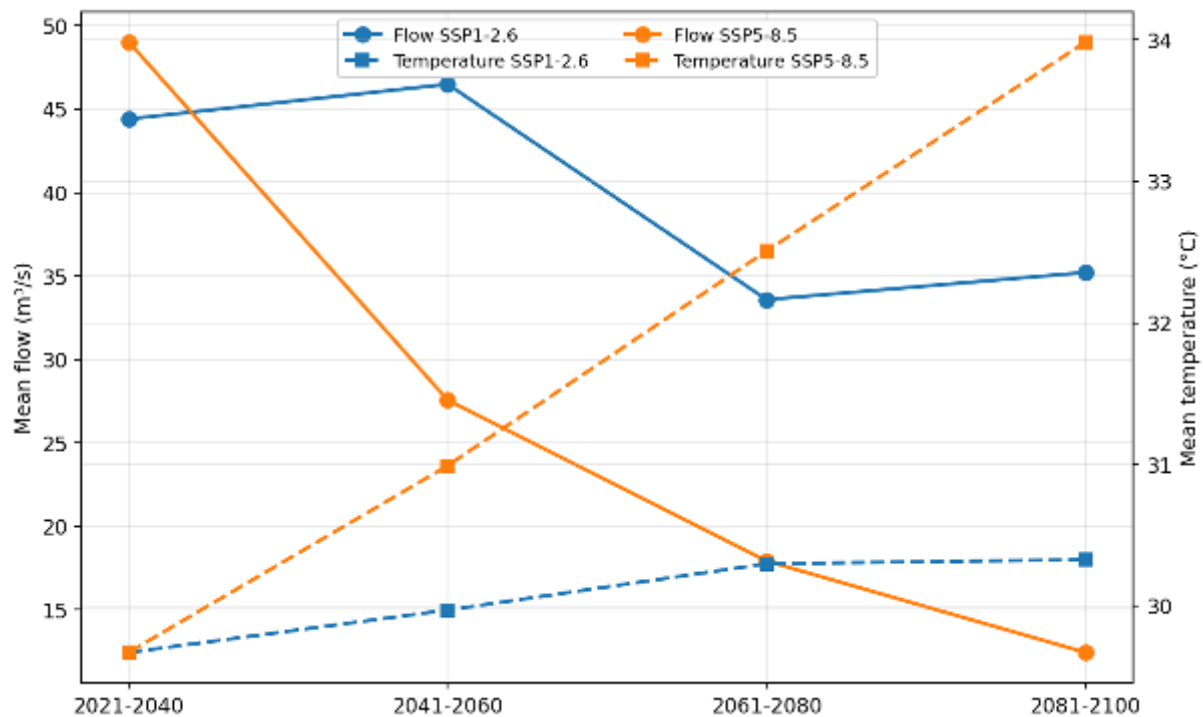


Figure 2. Joint evolution of mean streamflow and temperature across the four future periods.

Under SSP5-8.5, the thermal signal is much stronger. Mean temperature rises from 29.67 °C in 2021–2040 to 33.99 °C in 2081–2100 (+4.32 °C), while annual PET increases from 1,574.5 to 1,770.3 mm (+12.4%) (Figure 2). Precipitation varies between approximately 943 and 1,102 mm yr⁻¹ without a comparable monotonic decline. Nevertheless, mean streamflow collapses from 49.02 to 12.45 m³/s, a 74.6% decrease between these two periods. This divergence suggests that future hydrological drought cannot be interpreted from mean precipitation alone: increasing evaporative demand, seasonality, and nonlinear rainfall–runoff behavior play decisive roles.

3.2 Characterization Using SPI-3, SPEI-3, and SSI-3

The contrast between SPI-3 and SPEI-3 is one of the major findings. Under SSP5-8.5, mean SPI-3 remains between -0.39 and -0.11 depending on the period, whereas SPEI-3 reaches -1.59 in both 2041–2060 and 2081–2100 (Table 2). Incorporating PET therefore reveals a much stronger climatic deficit than precipitation alone would suggest. The same mechanism is evident under SSP1-2.6: in 2081–2100, SPI-3 is -0.22 whereas SPEI-3 reaches -1.14.

Table 2. Mean indices and frequency of hydrological drought classes.

Scenario	Period	SPI-3	SPEI-3	SSI-3	Drought (%)	Severe (%)	Extreme (%)
SSP1-2.6	2021-2040	-0.39	-1.16	-1.44	74.6	57.9	43.3
SSP1-2.6	2041-2060	-0.06	-1.01	-1.26	68.3	48.3	34.6
SSP1-2.6	2061-2080	-0.62	-1.54	-1.69	86.7	71.2	60.4
SSP1-2.6	2081-2100	-0.22	-1.14	-1.54	80.4	64.2	54.2
SSP5-8.5	2021-2040	-0.14	-0.85	-1.21	67.5	49.2	35.4
SSP5-8.5	2041-2060	-0.39	-1.59	-1.75	87.5	77.1	64.6
SSP5-8.5	2061-2080	-0.11	-1.42	-1.84	88.3	81.2	74.2
SSP5-8.5	2081-2100	-0.35	-1.59	-1.94	91.7	88.3	84.2

The hydrological response is even more severe. Under SSP1-2.6, mean SSI-3 is already below -1 in all periods: -1.44 in 2021–2040, -1.26 in 2041–2060, -1.69 in 2061–2080, and -1.54 in 2081–2100. The 2061–2080 period is the most unfavorable, with 86.7% of months in drought, 71.2% in severe drought, and 60.4% in extreme drought. The slight improvement in 2081–2100 does not return the system to historical conditions (Figure 3 and Figure 4).

Under SSP5-8.5, deterioration is progressive and very pronounced. Mean SSI-3 decreases from -1.21 in 2021–2040 to -1.75 in 2041–2060, -1.84 in 2061–2080, and -1.94 in 2081–2100. The frequency of drought months increases from 67.5% to 91.7%. Even more strikingly, extreme drought months rise from 35.4% to 84.2%, meaning that more than four out of five months in the final period fall below the $SSI-3 \leq -2$ threshold. The high-emissions pathway therefore leads to conditions approaching chronic hydrological drought relative to the 1985–2014 climatology.

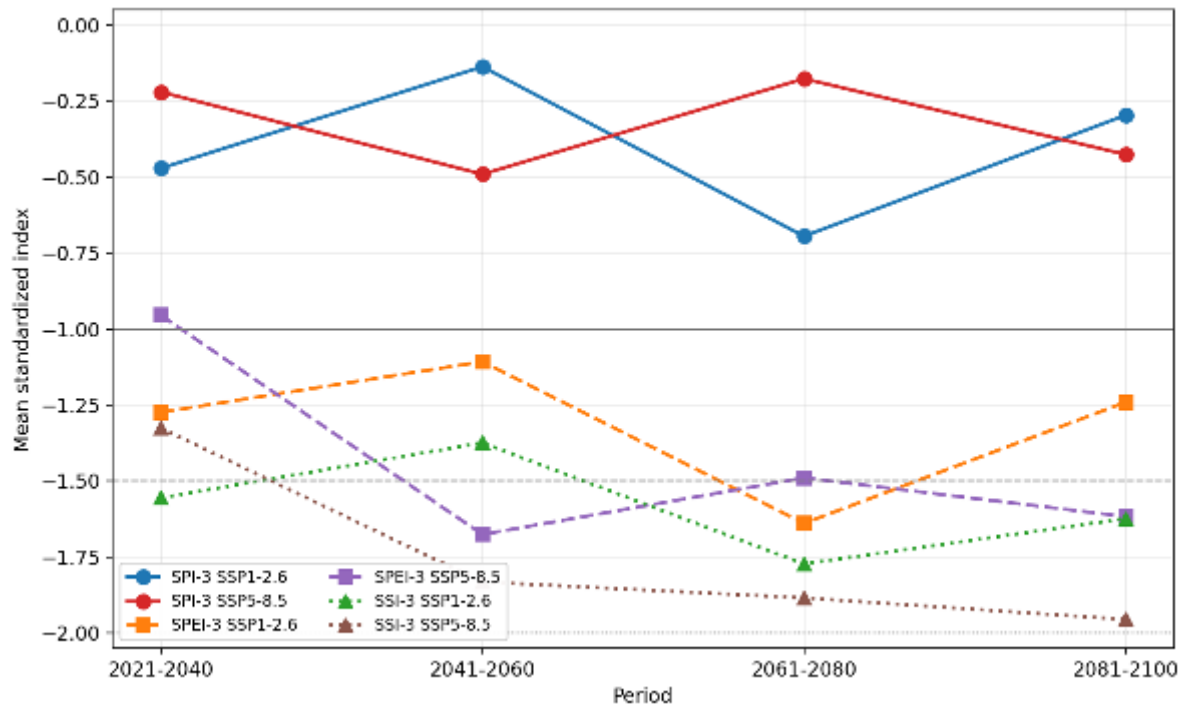


Figure 3. Evolution of mean SPI-3, SPEI-3, and SSI-3 by period and scenario.

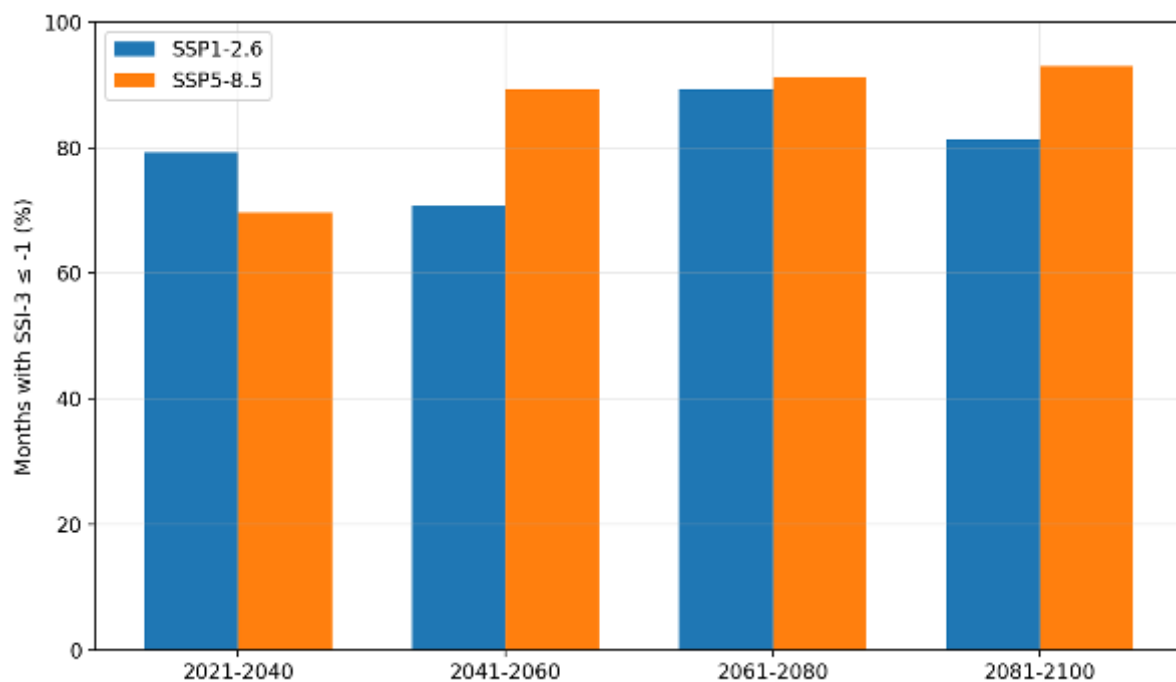


Figure 4. Proportion of months affected by hydrological drought (SSI-3 ≤ -1).

3.3 Duration and Severity of Drought Events

The event dataset identifies 68 episodes under SSP1-2.6 and 63 under SSP5-8.5. Event count alone is therefore not a sufficient indicator of worsening drought: under SSP5-8.5, episodes are longer on average (12.8 months versus 10.9 months) and more severe (mean cumulative severity of 25.3 versus 20.4) (Table 3 and Figure 5). Under SSP1-2.6, the longest continuous episode extends from June 2062 to April 2067, or 59 months. Under SSP5-8.5, several sequences

reach 35 months, including 2047–2050, 2051–2054, 2056–2059, and 2096–2099. The succession of long dry sequences under SSP5-8.5 confirms the tendency toward chronic drought indicated by monthly frequencies.

Table 3. Summary of hydrological drought events ($SSI-3 \leq -1$) over 2021–2100.

Scenario	Number	Mean duration (months)	Maximum duration	Mean severity	Longest event
SSP1-2.6	68	10.9	59	20.4	2062-06-01 to 2067-04-01
SSP5-8.5	63	12.8	35	25.3	2047-06-01 to 2050-04-01

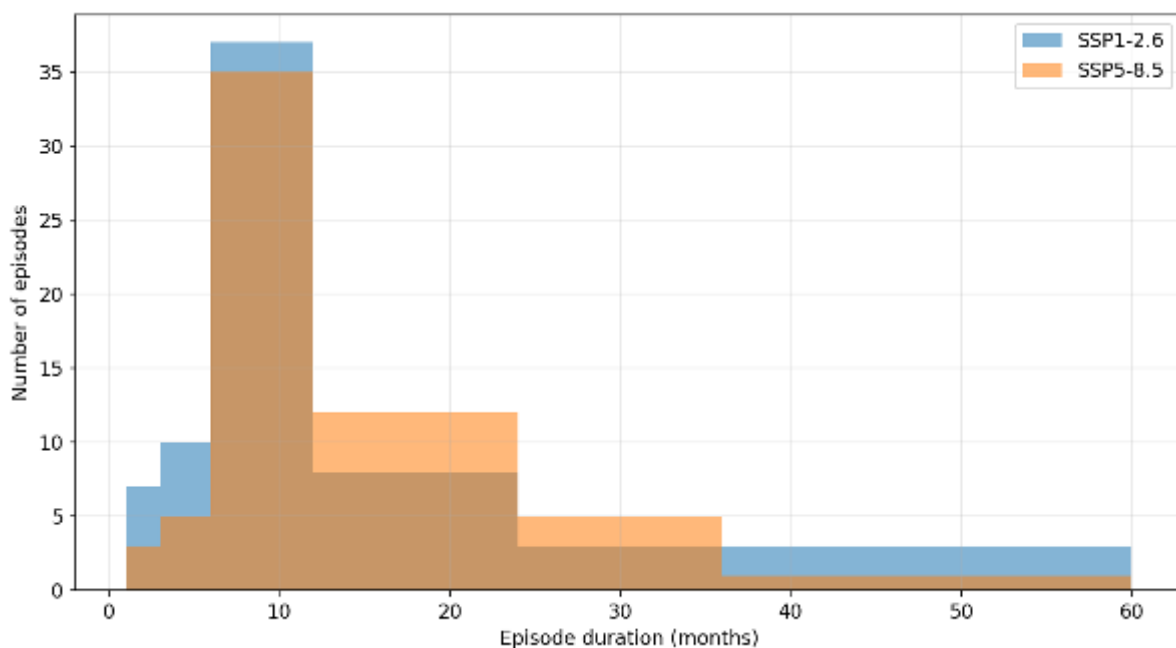


Figure 5. Distribution of hydrological drought-event duration.

3.4 Performance of SSI-3 Forecasting Models

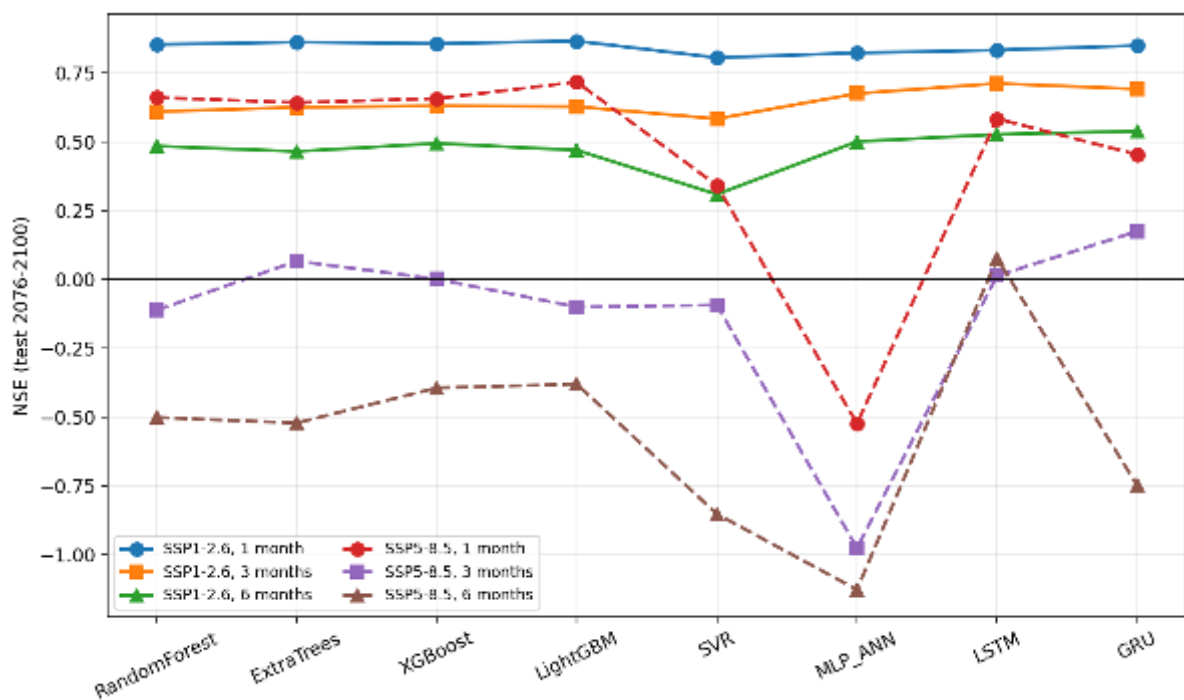
At a one-month lead time, boosting methods are particularly effective. LightGBM achieves the highest NSE under both scenarios: 0.867 under SSP1-2.6 ($KGE = 0.870$; $RMSE = 0.313$) and 0.718 under SSP5-8.5 ($KGE = 0.693$; $RMSE = 0.308$). Under SSP1-2.6, the other methods are also robust at one month: ExtraTrees reaches an NSE of 0.863, XGBoost 0.858, Random Forest 0.854, and GRU 0.850. The closeness of these scores indicates that recent system memory and lagged variables contain a strong predictive signal (Table 4 and Figure 6).

At three months under SSP1-2.6, LSTM becomes the best model with $NSE = 0.713$ and $F1 = 0.951$, followed by GRU ($NSE = 0.692$) and MLP/ANN (0.676). At six months, GRU retains the highest NSE (0.539), slightly ahead of LSTM (0.527) and MLP/ANN (0.501). This evolution is consistent with the value of recurrent architectures when the forecast horizon requires representation of longer temporal dependencies.

Table 4. Best model by scenario and forecast horizon according to test-set NSE.

Scenario	Lead time	Model	NSE	KGE	RMSE	MAE	F1
SSP1-2.6	1	LightGBM	0.867	0.870	0.313	0.183	0.960
SSP1-2.6	3	LSTM	0.713	0.671	0.459	0.314	0.951
SSP1-2.6	6	GRU	0.539	0.591	0.582	0.442	0.887
SSP5-8.5	1	LightGBM	0.718	0.693	0.308	0.179	0.978
SSP5-8.5	3	GRU	0.175	0.135	0.528	0.283	0.951
SSP5-8.5	6	LSTM	0.077	0.066	0.558	0.370	0.951

Under SSP5-8.5, forecasting difficulty increases sharply with lead time. At three months, the best NSE is only 0.175 for GRU; at six months, LSTM achieves 0.077. Several models even produce negative NSE values. Nevertheless, F1 scores remain close to 0.95 for the best recurrent networks. This contrast means that the models correctly recognize the drought state, which becomes very frequent, but have greater difficulty reproducing variations in SSI-3 intensity. In a context where more than 90% of months are dry by the end of the century, a high F1 score must therefore be interpreted jointly with NSE, KGE, and continuous error metrics.

**Figure 6. NSE of the eight models by scenario and forecast horizon.**

3.5 Variables Controlling Forecast Skill

At a one-month lead time, persistence in hydrological indices clearly dominates. Under SSP1-2.6, SSI3_lag0 accounts for 35.0% of LightGBM importance, SSI1_lag0 for 28.5%, and SSI1_lag1

for 7.7%. Under SSP5-8.5, SSI1_lag0 reaches 45.9%, SSI3_lag0 20.5%, followed by Discharge_lag2 at 6.7% and SSI1_lag1 at 5.8% (Figure 7). Short-term forecasting therefore depends primarily on the immediately preceding hydrological state.

As the forecast horizon increases, the role of climate forcing and streamflow memory becomes stronger. Under SSP1-2.6 at three months, Discharge_lag3 is the leading variable (25.2%), followed by SSI1_lag0 (12.7%) and Precipitation_lag6 (8.6%). At six months, Precipitation_lag2 becomes dominant (20.6%), ahead of Discharge_lag1 (10.4%). Under SSP5-8.5, precipitation lagged by two months is also the leading variable at six months (16.0%), followed by current streamflow (10.6%). These results highlight a transition from hydrological persistence at very short lead times to multi-timescale hydroclimatic control at seasonal horizons.

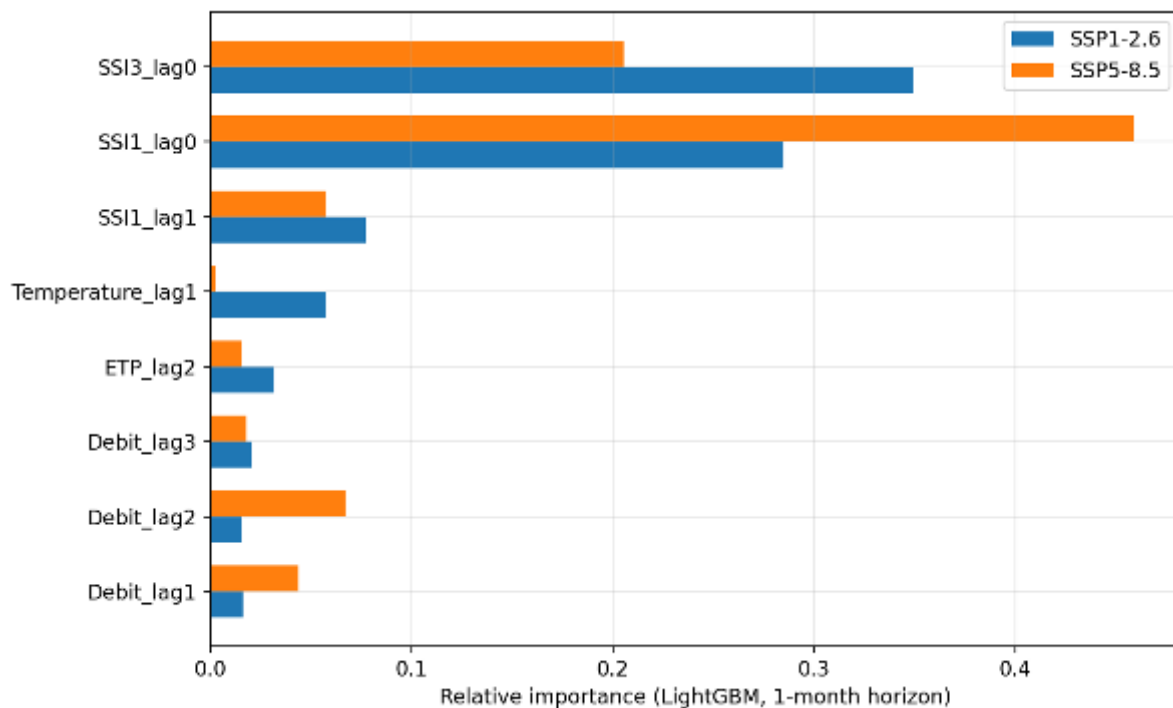


Figure 7. Main predictor importance values for the LightGBM model at a one-month lead time.

4. Discussion

The results reveal a marked divergence between meteorological, climatic, and hydrological drought responses in the Gambia River basin at Kédougou. While SPI-3 remains comparatively moderate across several future periods, SPEI-3 and especially SSI-3 show substantially stronger negative anomalies. This confirms that precipitation alone is insufficient to characterize future drought conditions, particularly under global warming, because drought results from interactions among precipitation deficits, atmospheric evaporative demand, catchment storage, and runoff processes (Mishra & Singh, 2010; Vicente-Serrano et al., 2010). The stronger deterioration of SPEI compared with SPI is therefore physically consistent with increasing temperature and potential evapotranspiration, while the

pronounced SSI response demonstrates the propagation and amplification of climatic anomalies through the hydrological system (Vicente-Serrano et al., 2012; Wang et al., 2022).

This amplification becomes particularly evident under SSP5-8.5. Between 2021–2040 and 2081–2100, mean temperature increases by 4.32 °C and PET by 12.4%, whereas mean streamflow decreases by approximately 74.6%, despite the absence of a comparable monotonic decline in precipitation. Such nonlinear rainfall–runoff behavior is consistent with the fundamental principles of catchment hydrology, whereby streamflow integrates precipitation, evapotranspiration, antecedent moisture conditions, storage, and basin characteristics (Beven, 2012). Comparable sensitivities have been reported for the Gambia basin, where rainfall–runoff simulations have demonstrated the importance of climatic forcing in controlling discharge variability (Faty et al., 2018; Faty et al., 2023). The projected contraction of flow also agrees with previous assessments indicating substantial climate-induced changes in the Senegal and Gambia river basins (Bodian et al., 2018; Rameshwaran et al., 2021).

The contrast between SSP1-2.6 and SSP5-8.5 further demonstrates the importance of emission pathways. Although SSP1-2.6 does not eliminate hydrological drought, it avoids the almost continuous deterioration projected under SSP5-8.5. This scenario dependence is consistent with the conceptual foundations of the Shared Socioeconomic Pathways and ScenarioMIP framework (Riahi et al., 2017; O'Neill et al., 2016). It is also consistent with recent basin-scale assessments showing that future hydroclimatic conditions in the Gambia basin are highly sensitive to climate forcing (Séne et al., 2024; Kah et al., 2025). The increasing frequency and persistence of extreme drought under SSP5-8.5 consequently suggest a transition from episodic drought toward quasi-chronic hydrological stress.

The forecasting results provide an additional methodological contribution. LightGBM achieves the highest performance at a one-month lead time, with NSE values of 0.867 under SSP1-2.6 and 0.718 under SSP5-8.5. This performance is consistent with the capacity of tree-based ensemble methods to capture nonlinear interactions among hydroclimatic predictors, as demonstrated more broadly for Random Forest approaches (Breiman, 2001). At longer horizons, however, LSTM and GRU become comparatively more effective, particularly under SSP1-2.6, confirming the relevance of recurrent architectures for learning temporal dependencies and hydrological memory (Zhao et al., 2024). Similar advantages of combining nonlinear machine-learning approaches with temporal information have been reported for long-term streamflow prediction (Abbasi et al., 2022).

Nevertheless, predictive skill deteriorates markedly at three- and six-month horizons under SSP5-8.5. This decline may reflect increasing hydroclimatic non-stationarity and the difficulty of reproducing SSI intensity when future conditions increasingly depart from the historical 1985–2014 reference distribution. Although F1 scores remain high, the low NSE values indicate that successful drought-state classification does not necessarily imply accurate prediction of drought magnitude. Model performance must therefore be evaluated using complementary metrics rather than a single criterion (Moriassi et al., 2007; Gupta et al., 2009).

Finally, several uncertainties remain. Future SSI values depend on simulated streamflow, meaning that hydrological-model uncertainty propagates into drought characterization.

Climate projections also contain uncertainties associated with model structure, emission scenarios, and regional climate representation (Abdulai & Chung, 2019; Almeida et al., 2024). Moreover, changes in land use can modify rainfall–runoff relationships and potentially reinforce or attenuate climate-driven hydrological responses (Ndiaye et al., 2024; Ba et al., 2024). These limitations support the future use of multi-GCM, multi-hydrological-model and ensemble forecasting frameworks. Nevertheless, the convergence of drought indices, streamflow projections, and artificial-intelligence models provides strong evidence that the upper Gambia River basin may face substantially greater hydrological drought pressure during the twenty-first century, particularly under high greenhouse-gas emissions.

The results suggest that future management of the upper Gambia River basin will need to address both declining mean streamflow and longer drought sequences. Under SSP5-8.5, the very high frequency of dry months by the end of the century implies potential constraints on water supply, agricultural uses, aquatic ecosystems, and river dilution capacity. Adaptation strategies should therefore strengthen hydroclimatic monitoring, conservation of recharge areas and riparian zones, water-use efficiency, seasonally appropriate storage, and mechanisms for transboundary information sharing.

From an early-warning perspective, the performance obtained at one month supports the development of an operational system combining LightGBM with a recurrent neural network. At three to six months, a probabilistic or ensemble approach would be preferable, particularly under SSP5-8.5. The strong importance of lagged SSI and streamflow indicates that an operational system must maintain continuous hydrometric observations; precipitation and PET become increasingly important for extending the forecast horizon.

5. CONCLUSION

This study provides an integrated assessment of future drought dynamics and predictability in the Gambia River basin at Kédougou by combining meteorological, climatic, and hydrological drought indices with machine-learning and deep-learning approaches. The joint analysis of SPI, SPEI, and SSI demonstrates that future drought cannot be adequately assessed from precipitation variability alone. The stronger deterioration of SPEI and, particularly, SSI highlights the critical influence of increasing evaporative demand and the amplification of climatic anomalies through the hydrological system.

The results reveal contrasting trajectories under SSP1-2.6 and SSP5-8.5. Although substantial hydrological drought persists under the low-emission pathway, SSP5-8.5 produces a pronounced and progressive deterioration of river-flow conditions. By 2081–2100, mean temperature reaches 33.99 °C, annual PET increases to 1,770.3 mm, and mean streamflow declines to 12.45 m³/s. Concurrently, mean SSI-3 falls to −1.94, while 91.7% of months are classified as drought and 84.2% as extreme drought. These changes indicate a potential transition from episodic drought toward quasi-chronic hydrological drought by the end of the century.

The forecasting experiment further demonstrates that model performance depends strongly on forecast horizon and climate scenario. LightGBM provides the best one-month forecasts,

reaching NSE values of 0.867 and 0.718 under SSP1-2.6 and SSP5-8.5, respectively. At longer lead times, LSTM and GRU become more competitive, particularly under SSP1-2.6. However, the sharp decline in continuous predictive performance at three and six months under SSP5-8.5 highlights the challenge posed by increasing hydroclimatic non-stationarity.

Overall, these findings demonstrate the value of combining standardized drought indicators with multimodel artificial-intelligence forecasting for drought early warning. They also emphasize the need to strengthen hydrometric monitoring, incorporate multi-model climate and hydrological uncertainties, and develop adaptive water-management strategies. Limiting greenhouse-gas emissions and improving anticipatory water governance appear complementary for reducing the risk of severe and persistent hydrological drought in the Gambia River basin.

Acknowledgements

The authors acknowledge the Directorate for Water Resources Management and Planning (DGPRES) of Senegal for providing streamflow data from the Kédougou gauging station. We also thank the Climate Hazards Group for the CHIRPS precipitation data, the European Centre for Medium-Range Weather Forecasts (ECMWF) for ERA5-Land temperature data, and the Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP) for providing climate projection data.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions (CRediT)

Ibrahima Faye: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Visualization, Writing - Original Draft. Cheikh FAYE: Conceptualization, Methodology, Investigation, Data Curation, Resources, Supervision, Project Administration, Validation, Writing - Review & Editing.

Ethics Approval and Consent to Participate

Ethical approval was not required for this study because the research uses publicly available environmental datasets (CHIRPS precipitation, ERA5-Land temperature, ISI-MIP3b climate projections) and secondary hydrometric data provided by the Directorate for Water Resources Management and Planning (DGPRES) of Senegal. No human subjects, animal experiments, or sensitive personal data were involved in this research.

Consent for Publication

The authors confirm that all necessary permissions and consents for publication were obtained, and the manuscript may be published accordingly.

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

Data availability: The streamflow data from the Kédougou gauging station used in this study were provided by the Directorate for Water Resources Management and Planning (DGPRE) of Senegal. These data are not publicly available due to institutional data-sharing policies but may be obtained from DGPRE upon reasonable request and with appropriate permission. CHIRPS precipitation data are publicly available at <https://www.chc.ucsb.edu/data/chirps> (Funk et al., 2015). ERA5-Land temperature data are available from the Copernicus Climate Change Service at <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land> (Muñoz-Sabater et al., 2021). ISIMIP3b climate projection data are available through the ISIMIP data portal at <https://www.isimip.org/gettingstarted/data-access/>.

Code availability: The Python code developed for this study, including the Random Forest rainfall–runoff model, standardized drought index calculations, and the eight machine-learning/deep-learning forecasting models, is available from the corresponding author upon reasonable request.

Declaration of Generative AI Use

The authors declare that generative artificial intelligence tools were not used for manuscript writing, data analysis, interpretation, or scientific conclusions.

ABBREVIATIONS AND NOTATION HARMONIZATION

Standardized Indices Notation

Index	Full Name	Accumulation Scale	Notation
SPI	Standardized Precipitation Index	3 months	SPI-3
SPEI	Standardized Precipitation Evapotranspiration Index	3 months	SPEI-3
SSI	Standardized Streamflow Index	3 months	SSI-3
SSI-1	Standardized Streamflow Index	1 month	SSI-1
SSI-6	Standardized Streamflow Index	6 months	SSI-6

Units and Variable Notation

Variable	Unit	Notation
Precipitation	mm yr ⁻¹ or mm month ⁻¹	P
Potential evapotranspiration	mm yr ⁻¹ or mm month ⁻¹	PET
Temperature	°C	T
Streamflow/Discharge	m ³ s ⁻¹	Q
Lagged variable	months	lag_n (e.g., lag_0, lag_1, lag_2)
Scenario Names		
Abbreviation	Full Name	
SSP1-2.6	Shared Socioeconomic Pathway 1-2.6 (low emissions)	
SSP5-8.5	Shared Socioeconomic Pathway 5-8.5 (high emissions)	

REFERENCES

- Abbasi, M., Dehban, H., Farokhnia, A., Roozbahani, R., & Bahreinimotlagh, M. (2022). Long-term streamflow prediction using hybrid SVR-ANN based on Bayesian model averaging. *Journal of Hydrologic Engineering*, 27(11), 05022018.
- Abdulai, P. J., & Chung, E. (2019). Uncertainty assessment in drought severities using multiple GCMs and the reliability ensemble averaging method. *Sustainability*, 11*(22), 6332. <https://doi.org/10.3390/su11226332>
- Ahmadalipour, A., Moradkhani, H., & Demirel, M. (2017). A comparative assessment of projected meteorological and hydrological droughts: Elucidating the role of temperature. *Journal of Hydrology*, 553*, 785–797. <https://doi.org/10.1016/j.jhydrol.2017.08.047>
- Ba, A., Bodian, A., Diop, L., & Sane, M. L. (2024). Land use/land cover changes and their impacts on hydrological processes in the Gambia River Basin. *Journal of Hydrology: Regional Studies*, 54*, 101885. <https://doi.org/10.1016/j.ejrh.2024.101885>
- Beven, K. J. (2012). *Rainfall-runoff modelling: The primer* (2nd ed.). Wiley-Blackwell.
- Bodian, A., Dezetter, A., Diop, L., Dème, A., Djaman, K., & Diop, A. (2018). Future climate change impacts on streamflows of two main West Africa river basins: Senegal and Gambia. *Hydrology*, 5*(1), 21. <https://doi.org/10.3390/hydrology5010021>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45*(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724–1734). Association for Computational Linguistics. <https://doi.org/10.3115/v1/D14-1179>

- de Almeida, L. P., Formiga-Johnsson, R. M., de Souza Filho, F. D. A., Estácio, Á. B. S., Porto, V. C., Nauditt, A., & Ribbe, L. (2024). Development of DRIP-drought representation index for CMIP climate model performance, application to Southeast Brazil. *Science of The Total Environment*, 954, 176443.
- Dione, O. (1996). *Évolution climatique récente et dynamique fluviale dans les hauts bassins des fleuves Sénégal et Gambie* [Doctoral dissertation, Université Jean Moulin Lyon 3].
- Diop, S. B., Ekolou, J., Trambly, Y., Dieppois, B., Grimaldi, S., Bodian, A., et al. (2025). Climate change impacts on floods in West Africa: New insight from two large-scale hydrological models. *Natural Hazards and Earth System Sciences*, *25*(9), 3161–3184. <https://doi.org/10.5194/nhess-25-3161-2025>
- Faty, B., Ali, A., Dacosta, H., Bodian, A., Diop, S., & Descroix, L. (2018). Assessment of satellite rainfall products for stream flow simulation in Gambia watershed. *African Journal of Environmental Science and Technology*, *12*(12), 501–513. <https://doi.org/10.5897/AJEST2018.2536>
- Faty, B., Sterk, G., Ali, A., Sy, S., Dacosta, H., Diop, S., & Descroix, L. (2023). Satellite-based rainfall estimates to simulate daily streamflow using a hydrological model over Gambia watershed. *Water Science*, *37*(1), 151–168. <https://doi.org/10.1080/23570008.2023.2234721>
- Faye, C., & Mendy, A. (2018). Climate variability and hydrological impacts in West Africa: The case of the Gambia River basin (Senegal). *Environmental and Water Sciences, Public Health*, *2*(1), 1–15.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, *2*, Article 150066. <https://doi.org/10.1038/sdata.2015.66>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Kah, M., Faye, C., Mbaye, M. L., Yalo, N., & Lischeid, G. (2025). Hydroclimatic trends and land use changes in the continental part of the Gambia River Basin: Implications for water resources. *Water*, *17*(14), 2075. <https://doi.org/10.3390/w17142075>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, *30*, 3146–3154.
- Liu, T., Si, Z., Liu, Y., et al. (2025). Runoff and drought responses to land use change and CMIP6 climate projections. *Water*, *17*(8), 1123. <https://doi.org/10.3390/w17081123>
- McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology* (pp. 179–184). American Meteorological Society.

- Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1–2), 202–216. <https://doi.org/10.1016/j.jhydrol.2010.07.012>
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900. <https://doi.org/10.13031/2013.23153>
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
- Ndiaye, P. M., Bodian, A., Diop, L., & Sane, M. L. (2024). Land use dynamics and their impacts on hydrological processes in the Gambia River Basin. *Journal of Hydrology: Regional Studies*, *53*, 101765. <https://doi.org/10.1016/j.ejrh.2024.101765>
- O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. <https://doi.org/10.5194/gmd-9-3461-2016>
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, *303*(1–4), 290–306. <https://doi.org/10.1016/j.jhydrol.2004.08.026>
- Rameshwaran, P., Bell, V. A., Davies, H. N., & Kay, A. L. (2021). How might climate change affect river flows across West Africa? *Climatic Change*, *169*(3–4), 21. <https://doi.org/10.1007/s10584-021-03256-0>
- Riahi, K., van Vuuren, D. P., Nakicenovic, N., Fricko, O., Krey, V., Gidden, M., et al. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. <https://doi.org/10.1016/j.gloenvcha.2016.05.009>
- Séne, S. M. K., Faye, C., & Pande, C. B. (2024). Assessment of current and future trends in water resources in the Gambia River Basin in a context of climate change. *Environmental Sciences Europe*, *36*(1), 32. <https://doi.org/10.1186/s12302-024-00856-w>
- Shukla, S., & Wood, A. W. (2008). Use of a standardized runoff index for characterizing hydrologic drought. *Geophysical Research Letters*, *35*(2), L02405. <https://doi.org/10.1029/2007GL032487>
- Song, Y., Shahid, S., & Chung, E. (2021). Differences in multi-model ensembles of CMIP5 and CMIP6 projections for future droughts in South Korea. *International Journal of Climatology*, *42*(5), 2688–2716. <https://doi.org/10.1002/joc.7386>
- Sow, A. A. (2007). *L'hydrologie du Sud-Est du Sénégal et de ses confins guinéo-maliens: Les bassins de la Gambie et de la Falémé* [Doctoral dissertation, Université Cheikh Anta Diop de Dakar].
- Traore, V., Sambou, S., Cisse, M., et al. (2014). Trends and shifts in time series of rainfall and runoff in the Gambia River Watershed. *International Journal of Environmental Protection and Policy*, *2*(4), 138–147. <https://doi.org/10.11648/j.ijepp.20140204.13>

- Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. <https://doi.org/10.1175/2009JCLI2909.1>
- Vicente-Serrano, S. M., López-Moreno, J. I., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C., & Morán-Tejeda, E. (2012). Accurate computation of a streamflow drought index. *Journal of Hydrologic Engineering*, *17*(2), 318–332. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000433](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000433)
- Wang, Y., et al. (2022). Drought propagation under global warming: Characteristics, approaches, processes, and controlling factors. *Science of the Total Environment*, *838*, 156021. <https://doi.org/10.1016/j.scitotenv.2022.156021>
- Zhao, X., Wang, H., Bai, M., Xu, Y., Dong, S., Rao, H., & Ming, W. (2024). A comprehensive review of methods for hydrological forecasting based on deep learning. *Water*, *16*(10), 1407. <https://doi.org/10.3390/w16101407>