Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning and Bias-Corrected CMIP6 for Improved Streamflow Prediction in the Data-Scarce Kulfo River, Ethiopia

Hailemariam Molla Ashagre

hailemariammolla25@gmail.com

Arba Minch Water Technology Institute, Arba Minch University

Demelash Wondimagegnehu Goshime

Arba Minch Water Technology Institute, Arba Minch University

Babur Tesfaye Yersaw

Arba Minch Water Technology Institute, Arba Minch University

Melaku Adugnaw Walle

Arba Minch Water Technology Institute, Arba Minch University

Research Article

Keywords: Streamflow projection, CMIP6, Bias-corrected, Machine learning, Climate change, Kulfo watershed

Posted Date: June 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10162255/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Discover Sustainability on September 18th, 2026. See the published version at https://doi.org/10.1007/s43621-026-04724-2.

Machine Learning and Bias-Corrected CMIP6 for Improved Streamflow Prediction in the Data-Scarce Kulfo River, Ethiopia

Hailemariam Molla Ashagre*[a], Demelash Wondimagegnehu Goshime[b], Babur Tesfaye Yersaw[b], Melaku Adugnaw Walle [b]

[a] Faculty of Hydraulic and Water Resources Engineering, Arba Minch Water Technology Institute, Arba Minch University, Ethiopia

[b] Faculty of Water Resources and Irrigation Engineering, Arba Minch Water Technology Institute, Arba Minch University, Ethiopia

Corresponding author*: hailemariammolla25@gmail.com

Highlights

- Better-performing CMIP6 models: CNRM-ESM2-1 for precipitation and minimum temperature, and CNRM-CM6-1 for maximum temperature after bias correction.
- Random Forest bias correction reduced precipitation RMSE by 92%.
- Random Forest achieved strong performance (NSE = 0.999 training; 0.806 validation).
- Climate projections show rising temperatures and declining streamflow under all SSP scenarios.
- Conformal prediction quantified uncertainty using 95% prediction intervals.

Abstract

Climate change is increasingly altering hydrological processes and water availability, particularly in data-scarce regions such as Ethiopia's Rift Valley basins. This study integrated bias-corrected Coupled Model Intercomparison Project Phase 6 (CMIP6) climate projections with machine learning approaches to evaluate future streamflow dynamics in the Kulfo watershed, southern Ethiopia. CMIP6 models under shared socio-economic pathways namelySSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios were evaluated using precipitation and temperature observations from 1986–2014. Bias correction was applied using statistical and machine learning approaches to improve climate model reliability before hydrological application, while model performance was evaluated using Nash–Sutcliffe Efficiency (NSE),

coefficient of determination (R^2), Root Mean Square Error (RMSE), and Kling–Gupta Efficiency (KGE). Results showed that bias correction significantly improved CMIP6 model performance, reducing precipitation RMSE by averagely 83% and substantially improving temperature simulations. Among the developed models, Random Forest (RF) demonstrated the best predictive capability with NSE values of 0.999 during calibration and 0.806 during validation. Future climate projections indicate increasing maximum and minimum temperatures under all SSP scenarios, while annual streamflow is projected to decline progressively, reaching reductions of approximately 8.5%, 13.9%, and 20.7% under SSP1-2.6, SSP2-4.5, and SSP5-8.5, respectively. Seasonal analysis revealed substantial reductions in summer discharge and moderate increases in winter flow, indicating a redistribution of hydrological seasonality. Conformal prediction analysis further showed increasing uncertainty in future streamflow projections, particularly under SSP5-8.5, while non-stationary flood frequency analysis using a time-varying Generalized Extreme Value (GEV) model demonstrated significant future changes in flood characteristics under climate change conditions. The findings reveal the vulnerability of the Kulfo watershed to future hydro-climatic changes and emphasize integrating machine learning, uncertainty analysis, and non-stationary hydrological approaches into climate-resilient water resource planning to support sustainable watershed management, irrigation development, flood risk reduction, and climate adaptation in Ethiopia and other data-scarce tropical catchments.

Keywords: Streamflow projection; CMIP6; Bias-corrected; Machine learning; Climate change; Kulfo watershed

1 Introduction

Climate change has emerged as one of the most critical challenges affecting hydrological processes and water resources availability worldwide (Calvin et al. 2023a). The Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report highlights that global surface temperatures have increased by approximately 1.1°C since the late nineteenth century, with profound implications for the hydrological cycle, including altered precipitation patterns, increased evaporation rates, and intensified extremes such as floods and droughts (Intergovernmental Panel On Climate Change (Ipcc) 2023a). East Africa, in particular, has been

identified as a climate change hotspot, characterized by heightened vulnerability due to limited adaptive capacity, rapid population growth, and heavy dependence on rain-fed agriculture (Jin et al. 2022).

Ethiopia's water resources are exceptionally susceptible to climate variability and change (Teku 2025). The country has experienced recurrent multi-season droughts and devastating floods over recent decades, causing substantial economic losses and humanitarian crises (Bezu 2020). The Kulfo watershed, located in southern Ethiopia's Rift Valley region near Arba Minch, exemplifies these challenges. The watershed supports agricultural livelihoods, ecosystem services including the Nech Sar National Park, and downstream communities along the shores of Lake Abaya, yet lacks comprehensive hydrological assessments that integrate state-of-the-art climate projections with advanced modeling techniques (Truneh et al. 2022).

Traditional hydrological modeling approaches often struggle to adequately represent the complex, non-linear relationships between climate variables and stream flow responses, particularly in data-sparse regions (Näschen et al. 2018). Process-based models such as the Soil and Water Assessment Tool (SWAT) require extensive parameterization and calibration data that are frequently unavailable in developing country contexts (Kadlag et al. 2025). Furthermore, Global Climate Models (GCMs) from the Coupled Model Intercomparison Project Phase 6 (CMIP6) exhibit systematic biases when applied at local scales, necessitating rigorous bias correction and downscaling before hydrological applications (Eyring et al. 2016; Rohling et al. 2022).

Recent advances in machine learning offer promising alternatives for hydrological modeling, capable of capturing non-linear dynamics without explicit representation of physical processes (Kumar et al. 2023; Gacu et al. 2025). However, limited studies have systematically evaluated multiple machine learning algorithms for stream flow projection in Ethiopian basins using the latest CMIP6 climate scenarios (Abdulahi et al. 2022; Gebresellase et al. 2025). The integration of bias-corrected CMIP6 projections with machine learning-based hydrological models represents an innovative approach to address data limitations while providing robust future stream flow estimates for the Kulfo watershed in southern Ethiopia.

This research contributes to the scientific literature by providing the first comprehensive assessment of CMIP6 model performance in the Kulfo watershed, demonstrating the superiority of ensemble machine learning approaches for streamflow projection in data-scarce Ethiopian catchments, integrating multiple CMIP6 models and scenarios to address projection uncertainty, and offering practical insights for water resource planning and policy development in the Arba Minch region.

2 Material and Methods

2.1 Study Area

2.1.1 Geographic Location

The Kulfo watershed is situated in the Southern Ethiopia, approximately 500 km south of Addis Ababa and adjacent to the town of Arba Minch. The watershed lies between 5.9°-6.2° N latitude and 37.5°-37.8° E longitude, draining from the western slopes of the Rift Valley escarpment into the eastern shore of Lake Abaya, Ethiopia's largest Rift Valley Lake (Lemma et al., 2023; Ayenew et al., 2023). The river originates from the Gughe Highlands at elevations exceeding 3,000 meters above sea level and flows approximately 45 km through the Arba Minch plains before its deltaic entry into Lake Abaya.

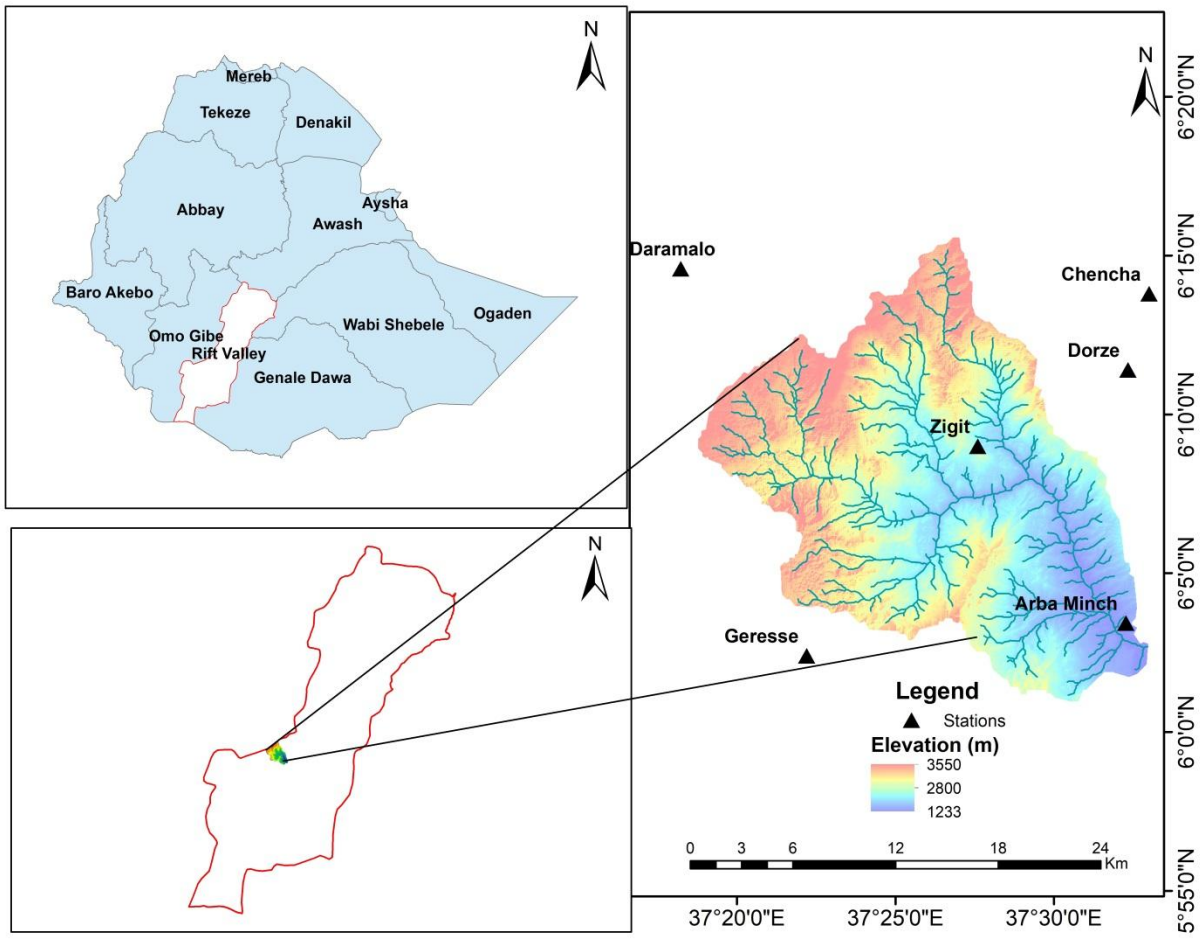


Fig. 1 Location map of Kulfo Watershed, southern Ethiopia, showing elevation zones, River networks

2.1.2 Climate and Hydrology

The Kulfo watershed experiences a bimodal rainfall pattern characteristic of equatorial East Africa. The primary rainy season (Belg) occurs from March to May, contributing approximately 35% of annual precipitation, while the secondary rainy season (Kiremt) extends from June to September, contributing 50% of annual rainfall (Orke and Li 2021). The dry seasons occur from October to February, with minimal precipitation. Mean monthly temperatures range from 20°C in July-August to 25°C in February-March.

The Kulfo River is perennial, sustained by groundwater contributions from the highland aquifers during dry periods (Ayalew et al. 2023a). Mean annual discharge at the Abaya gauging station is approximately 8.5 m^3/s, with peak flows exceeding 80 m^3/s during August-September

flood events. The river provides critical water supplies for Arba Minch town (population ~150,000), supports approximately 8,500 ha of irrigated agriculture, and maintains the ecologically significant Kulfo wetland at its terminus.

2.2 Observed Hydro-meteorological Data

2.2.1 Precipitation Data

Daily precipitation records (1986-2014) were obtained from the Ethiopian Meteorology Institute (EMI) for the Arba Minch meteorological station (6.03°N, 37.55°E, and elevation 1,250 m). The dataset underwent rigorous quality control procedures including: (1) identification of missing values (<3% of records), (2) detection of outliers using interquartile range methodology, (3) homogeneity testing using the Pettitt test and Standard Normal Homogeneity Test (SNHT), and (4) gap-filling using correlation with neighboring stations (Chencha and Gidole) where applicable (Yimer et al. 2023).

2.2.2 Temperature Data

Daily maximum and minimum temperature records (1986-2014) were obtained from the same Arba Minch station. Quality control procedures identical to precipitation were applied. Temperature records exhibited no significant missing data (<1%) and demonstrated strong temporal consistency.

2.2.3 Stream flow Data

Daily stream flow records (1980–2022) for the Kulfo River were obtained from the Ministry of Water and Energy (MoWE) gauging station located at the Abaya bridge (6.01°N, 37.58°E; elevation 1,180 m). Data quality assessment indicated that records were complete for approximately 92% of the study period, with gaps primarily occurring during 1991–1993, coinciding with a period of civil unrest. The uncertainty associated with the rating curves was estimated at ±8% for medium flows and ±15% for high flows, while no significant trends or shifts were observed in the stage-discharge relationship.

2.3 CMIP6 Climate Model Data

2.3.1 Model Selection Criteria

Sixteen(16) CMIP6 global climate models were selected based on: (1) availability of daily precipitation, maximum temperature, and minimum temperature variables, (2) availability of historical simulations (1986–2014) and future projections (2015–2100) under SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios, (3) spatial resolution suitable for East African applications, and (4) documented performance in previous East African studies (Gusain et al. 2020a; Almazroui et al. 2020; Akinsanola et al. 2021).

Table 1 Selected CMIP6 Models and Specifications

Model	Country / Institute	Resolution	Citation
ACCESS-CM2	Australia – CSIRO & BoM	1.25° × 1.875°	(Dix et al. 2020)
CanESM5	Canada – Environment & Climate Change Canada	2.8° × 2.8°	(Swart et al. 2019)
CMCC-ESM2	Italy – CMCC	1.25° × 1.0°	(Cherchi et al. 2019)
CNRM-CM6-1	France – CNRM/CERFACS	1.4° × 1.4°	(Séférian et al. 2019)
CNRM-ESM2-1	France – CNRM/CERFACS	1.4° × 1.4°	(Séférian et al. 2019)
EC-Earth3-Veg-LR	Europe – EC-Earth Consortium	1.0° × 1.0°	(Döscher et al., 2022)
GFDL-ESM4	USA – NOAA GFDL	1.0° × 1.0°	(Dunne et al. 2020)
INM-CM4-8	Russia - Institute of Numerical Mathematics	2.0° × 1.5°	(Volodin et al., 2017)
INM-CM5-0	Russia – Institute of Numerical Mathematics	2.0° × 1.5°	(Volodin et al., 2017)
KACE-1-0-G	South Korea – NIMS/KMA	1.25° × 1.25°	(Byun et al. 2019)
MIROC6	Japan – JAMSTEC, AORI, NIES	1.4° × 1.4°	(Tatebe et al. 2019)
MIROC-ES2L	Japan – NIES, JAMSTEC	2.8° × 2.8°	(Hajima et al. 2020)
MPI-ESM1-2-LR	Germany – Max Planck Institute for Meteorology	1.9° × 1.9°	(Mauritsen et al. 2019)
MRI-ESM2-0	Japan – Meteorological Research Institute	1.125° × 1.125°	(Yukimoto et al. 2019)
NESM3	China – Nanjing University of Information Science and Technology	1.25° × 1.25°	(Cao et al. 2018)
NorESM2-MM	Norway – Norwegian Climate Centre	1.25° × 0.94°	(Seland et al., 2020)

2.3.2 Scenario Development

Three Shared Socioeconomic Pathway (SSP) scenarios were selected to represent a range of possible future socioeconomic and climate trajectories (Riahi et al. 2017; Tebaldi et al. 2020): SSP1-2.6, a low-emissions scenario consistent with the goals of the Paris Agreement and reflecting a sustainability pathway; SSP2-4.5, an intermediate emissions scenario representing a middle-of-the-road development pathway; and SSP5-8.5, a high-emissions scenario

characterized by fossil-fueled development. Together, these scenarios provide comprehensive coverage of future uncertainty and are consistent with guidance from the Intergovernmental Panel on Climate Change for climate impact assessments (Tebaldi et al. 2020; Calvin et al. 2023b).

2.4 Bias correction using Machine Learning

The best-performing bias correction methods were identified from both statistical and machine learning approaches based on predefined evaluation metrics. Following this selection, raw outputs from CMIP6 were corrected to reduce systematic biases relative to local observations. These biases arise from coarse spatial resolution, parameterization uncertainties, and the inadequate representation of local topographic effects (Teutschbein and Seibert 2012a). To address these limitations, the selected statistical and machine learning-based bias correction techniques were applied to the Kulfo watershed, ensuring improved reliability and consistency of the corrected climate projections.

Bias correction was calibrated using the 1986–2006 periods and validated against 2007-2014 observations. Performance metrics included percent bias (PBIAS), root mean square error (RMSE), and correlation coefficient.

2.5 Machine Learning Model Development

2.5.1 Conceptual Framework

The rainfall–runoff relationship was conceptualized as a function of current and antecedent hydro-climatic variables, expressed as:

$$Q_t = f(P_t, P_{t-1}, P_{t-2}, Tmax_t, Tmin_t, Q_{t-1})$$

Where; Qt represents streamflow at time t, Pt denotes precipitation at time t, and P_{t-1} and $P_{t-2}$are antecedent precipitation values at 1- and 2-day lags, respectively. T_{maxt} and T_{mint} correspond to the maximum and minimum temperatures at time t, while Qt−1represents antecedent streamflow with a 1-day lag. This formulation captures both the immediate and delayed hydrological responses of the watershed to climatic inputs.

2.5.2 Data Preprocessing

Data were partitioned into a training period 1986-2006, 70% of the dataset and a testing period 2007-2014, 30%. All input variables were normalized to zero mean and unit standard deviation to ensure that each variable contributed equitably across the modeling algorithms

2.5.3. Application of Machine Learning Models in Hydrological Forecasting

The machine learning and deep learning models represent a combination of ensemble, neural network, instance-based, and kernel-based approaches that are increasingly applied in hydrology, climate prediction, and environmental modeling studies. Ensemble learning techniques such as Random Forest (RF), Bagging, Boosting (GBM), AdaBoost, CatBoost, and Stacking are widely recognized for their ability to improve prediction accuracy and reduce uncertainty in complex datasets. RF is particularly effective in handling nonlinear hydrological relationships and high-dimensional climate variables through multiple decision-tree aggregation, while boosting algorithms iteratively improve weak learners to minimize prediction errors. CatBoost has recently gained significant attention because of its efficient handling of categorical variables and reduced overfitting characteristics. Stacking further enhances predictive performance by combining multiple base learners using a meta-model. Recent studies demonstrated that ensemble machine learning methods significantly improve stream flow forecasting, rainfall–runoff simulation, and water resources prediction under varying climatic conditions (Li et al. 2024; Gelete et al. 2025).

In addition to ensemble methods, the study employed Long Short-Term Memory (LSTM), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) models to capture temporal and nonlinear patterns in hydrological datasets. Long Short-Term Memory models are particularly effective for time-series forecasting because they retain long-term dependencies through memory cells and gating mechanisms, making them highly suitable for rainfall, streamflow, and climate prediction. Recent hydrological studies have shown that LSTM and hybrid deep learning frameworks outperform many traditional statistical and machine learning approaches in sequential prediction tasks. KNN is a distance-based algorithm that estimates outputs using neighboring observations with similar characteristics, whereas SVM uses kernel-based optimization to model nonlinear relationships with strong generalization capability. The

integration of these diverse algorithms provides a robust framework for improving prediction reliability, minimizing uncertainty, and enhancing hydrological forecasting accuracy (De La Fuente et al. 2025; Weißenborn 2025).

Table 2 Optimized Hyper parameters for Machine Learning Models

Model	Hyper parameter	Tested Range	Selected Value
RF (Random Forest)	Number of trees	100, 200, 500, 1000	500
	Max depth	10, 20, 30, None	20
	Min samples split	2, 5, 10	5
	Min samples leaf	1, 2, 4	2
LSTM (Long Short-Term Memory)	Number of layers	1, 2, 3	2
	Units per layer	32, 64, 128	64
	Learning rate	0.0001, 0.001, 0.01	0.001
	Batch size	32, 64, 128	64
	Epochs	50, 100, 200	100
	Dropout rate	0.0, 0.2, 0.5	0.2
Bagging	Base estimator	Decision Tree	Decision Tree
	Number of estimators	10, 50, 100	50
	Max samples	0.5, 0.7, 1.0	0.7
Boosting (GBM)	Number of estimators	50, 100, 200	100
	Learning rate	0.01, 0.1, 0.2	0.1
	Max depth	3, 5, 7	5
AdaBoost	Number of estimators	50, 100, 200	100
	Learning rate	0.01, 0.1, 1.0	0.1
CatBoost	Iterations	100, 500, 1000	500
	Learning rate	0.01, 0.05, 0.1	0.05
	Depth	4, 6, 8	6
KNN	Number of neighbors (K)	3, 5, 7, 9	5
	Weights	uniform, distance	distance
	Metric	euclidean, manhattan	euclidean
SVM	C	0.1, 1, 10, 100	10
	Kernel	linear, rbf, poly	rbf
	Gamma	0.001, 0.01, 0.1	0.01

2.5.3 Model Performance Evaluation

The performance of the developed rainfall–runoff models was evaluated using multiple statistical metrics to assess accuracy, reliability, and predictive skill. Key evaluation criteria included the Coefficient of Determination (R^2) to measure the proportion of variance explained, Nash–Sutcliffe Efficiency (NSE) for overall model predictive performance, Root Mean Square Error (RMSE) for quantifying average prediction error magnitude, and Mean Absolute Error (MAE) to

assess bias in predictions. Model evaluation was conducted separately for the training, validation, and testing periods to ensure robustness and generalization capability across different temporal subsets of the data.

2.6 Current and Future Stream flow Projection Framework

The current stream flow used for the study was 1986-2014 and the projection framework followed a sequential approach. CMIP6 precipitation and temperature outputs for 2015-2100 were bias-corrected under three SSP scenarios. These corrected climate variables were then input into the best-performing machine learning model to simulate daily stream flow for each CMIP6 model and scenario. The resulting daily stream flow was aggregated to monthly, seasonal, and annual timescales. Changes were computed relative to the baseline period (1986-2014), and uncertainty ranges were quantified across the CMIP6 ensemble to capture variability among models and scenarios.

2.7 Probabilistic projections with conformal prediction

CP (Conformal prediction) (Shafer & Vovk, 2008) was integrated with the Random Forest (RF) model to quantify uncertainty in future river discharge projections under SSP126, SSP245, and SSP585 scenarios. The method generated 95% prediction intervals (PI) using residual-based non-conformity scores derived from the calibration period, providing statistically valid upper and lower bounds for projected streamflow. The probabilistic projections enabled assessment of the reliability and spread of future discharge estimates while accounting for uncertainty associated with climate variability and model prediction errors. The interval widths generally increased toward the end of the century, indicating greater uncertainty in long-term projections. In addition, asymmetric interval growth was evaluated by comparing the expansion rates of upper and lower PI bounds to determine whether uncertainty was more strongly associated with potential increases in extreme flows or reductions in low-flow conditions. Similar studies have demonstrated the usefulness of conformal prediction for improving uncertainty quantification and reliability in hydrological forecasting applications (Sesia and Candès 2020).

2.8 Non-stationary flood frequency analysis

A non-stationary flood frequency analysis was performed using annual maximum discharge series to evaluate the influence of climate change on future flood characteristics. A time-varying generalized extreme value (GEV) model was fitted to annual maximum flows, with the location parameter expressed as $\mu(t) = \mu_0 + \mu_1 \cdot \Delta T(t)$, where μ_0 is the baseline location parameter, μ_1 represents the sensitivity of flood magnitude to temperature change, and $\Delta T(t)$ is the projected temperature anomaly at time t. This non-stationary framework allowed flood probabilities and return periods to evolve over time rather than remain constant, thereby capturing the effects of climatic change on hydrological extremes. Future flood quantiles and recurrence intervals were re-estimated for SSP126, SSP245, and SSP585 scenarios to assess changes in flood risk throughout the twenty-first century. Incorporating non-stationarity into flood frequency analysis is increasingly recommended because climate change can alter the magnitude and frequency of extreme hydrological events, making stationary assumptions less reliable for long-term water resources planning and flood risk management (Twinomuhangi et al. 2025).

3 Results and Discussion

3.1 Performance evaluation of CMIP6 Models

3.1.1 Precipitation (pr)

The performance of multiple climate models in simulating precipitation before and after bias correction over the period 1986–2014 is shown in Figure 2. In the pre-correction stage (Figure a), the model outputs exhibit substantial variability and deviate significantly from the observed data, with some models overestimating precipitation while others underestimate it, indicating the presence of systematic biases and weak agreement with observed rainfall patterns. Such discrepancies are common in raw climate model outputs due to limitations in representing local-scale processes and regional climate variability. However, after applying bias correction (Figure b), the model simulations show a marked improvement, with most corrected outputs closely aligning with the observed precipitation. The spread among models is significantly reduced, and they effectively capture the temporal variability and trends of the observed data. The highlighted model (CNRM-ESM2-1) demonstrates particularly strong agreement with observations. Overall, these results confirm that bias correction is an essential step in enhancing the accuracy and reliability of climate model precipitation simulations, making them more suitable for hydrological and climate impact studies (Luo et al. 2025a).

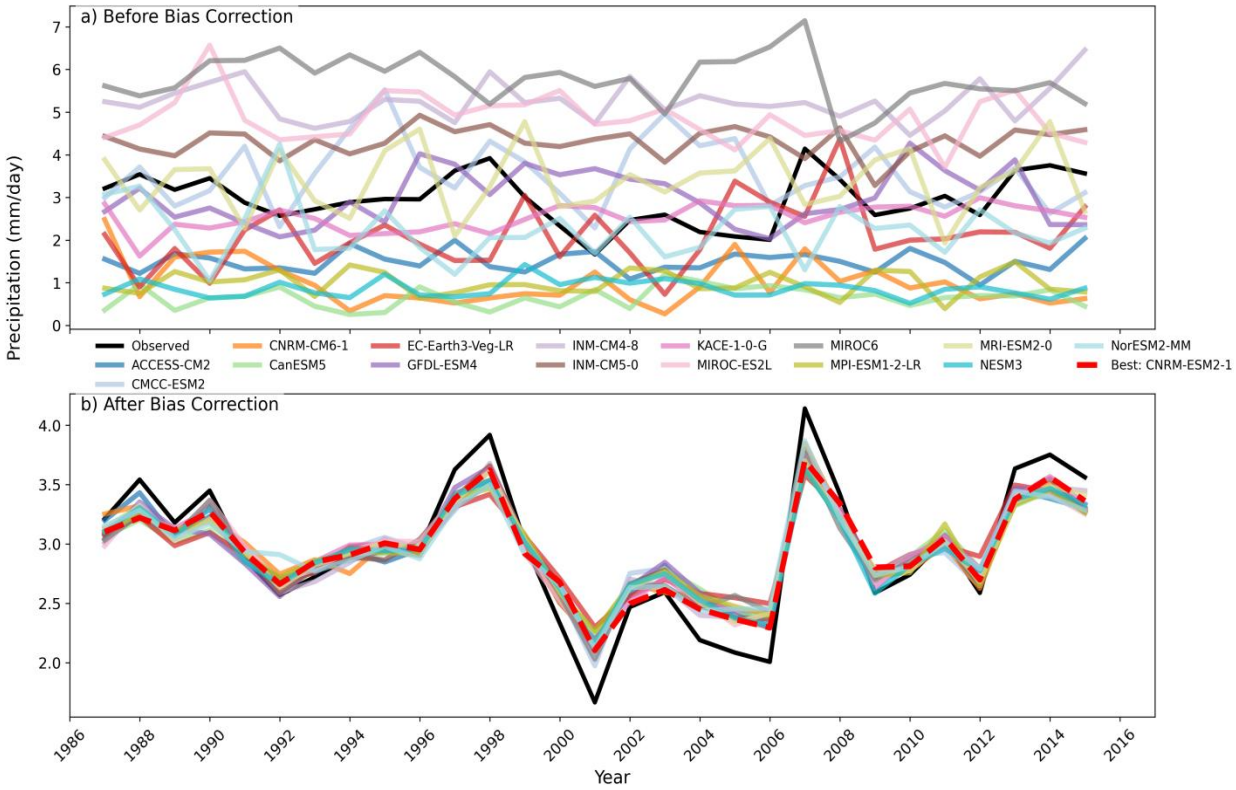


Fig. 2**:** Annual precipitation (mm/day) time series from 1986–2014 comparing observations with 16 CMIP6 models: (a) before bias correction, (b) after Random Forest correction

The performance of various climate models before and after bias correction using RMSE, correlation coefficient (R), percent bias (PBIAS), and Kling-Gupta Efficiency (KGE), along with their improvements (Table 5). Prior to bias correction, most models show poor performance, with high RMSE values (up to 2.95mm), weak or negative correlations, large biases (both overestimation and underestimation), and negative KGE values, indicating low agreement with observed precipitation. Such limitations are commonly reported in raw climate model outputs due to structural uncertainties and scale mismatches **(Teutschbein and Seibert 2012b; Maraun 2016; Vogel et al. 2023)**. After bias correction, all models exhibit significant improvement, as RMSE values decrease substantially to about 0.21-0.29 mm, correlation coefficients increase to very high levels (0.97-0.99), PBIAS values approach zero, and KGE values improve to positive ranges (approximately 0.54-0.67 %), reflecting better overall model performance. The absolute reduction in RMSE is considerable (-0.64 to -2.72 mm), with relative improvements exceeding 74% for all models and surpassing 90% for some, such as MIROC6 and INM-CM4-8. Notably,

models like CNRM-ESM2-1 and CNRM-CM6-1 show excellent post-correction performance with low error, strong correlation, and minimal bias. Overall, the results clearly indicate that bias correction greatly enhances the accuracy, consistency, and reliability of climate model precipitation simulations, making them more suitable for hydrological and climate impact assessments **(Dhawan et al. 2024; Luo et al. 2025b)**.

Table 3 Raw CMIP6 Performance -Precipitation (1986-2014)

Models	Before bias				After bias				Absolute Improvement (mm)	Relative Improvement (%)
	RMSE (mm)	R	PBIAS (%)	KGE	RMSE (mm)	R	PBIAS (%)	KGE		
CNRM-ESM2-1	2.10	0.02	-66.43	-0.21	0.21	0.99	-0.07	0.67	-1.89	-89.99
CNRM-CM6-1	2.11	-0.01	-66.22	-0.21	0.22	0.98	0.01	0.67	-1.90	-89.73
MIROC-ES2L	2.04	0.11	64.04	-0.10	0.22	0.99	-0.01	0.64	-1.82	-89.04
KACE-1-0-G	0.89	-0.47	-14.44	-0.56	0.22	0.99	0.41	0.65	-0.67	-75.26
INM-CM4-8	2.39	0.22	77.66	-0.13	0.22	0.99	-0.22	0.66	-2.17	-90.79
INM-CM5-0	1.49	0.16	45.70	-0.06	0.22	0.99	-0.19	0.64	-1.27	-85.02
GFDL-ESM4	0.86	0.00	1.48	0.00	0.22	0.98	0.10	0.66	-0.64	-74.30
ACCESS-CM2	1.59	0.13	-49.40	-0.15	0.23	0.99	-0.07	0.64	-1.36	-85.71
MRI-ESM2-0	1.12	-0.13	15.01	-0.17	0.23	0.99	0.11	0.63	-0.89	-79.72
NESM3	2.20	-0.17	-70.83	-0.52	0.23	0.99	-0.26	0.62	-1.96	-89.43
MIROC6	2.95	-0.11	95.23	-0.47	0.23	0.99	-0.19	0.64	-2.72	-92.12
CMCC-ESM2	1.15	-0.02	20.46	-0.08	0.24	0.99	-0.07	0.63	-0.92	-79.56
MPI-ESM1-2-LR	2.08	-0.17	-66.39	-0.45	0.26	0.99	0.08	0.57	-1.82	-87.33
CanESM5	2.40	-0.25	-77.80	-0.59	0.27	0.99	-0.05	0.57	-2.13	-88.79
NorESM2-MM	1.22	-0.24	-24.16	-0.26	0.24	0.98	-0.05	0.62	-0.97	-79.94
EC-Earth3-Veg-LR	1.33	-0.12	-28.97	-0.19	0.29	0.97	0.30	0.54	-1.03	-77.84

3.1.2 Maximum temperature (Tmax)

The figure (a) and (B) compares observed and modeled annual maximum temperature (1986–2014). Before bias correction, large spread and systematic over- or underestimation occurs. After correction, closer agreement with observations and reduced variability are evident, showing improved accuracy and reliability of climate model outputs.

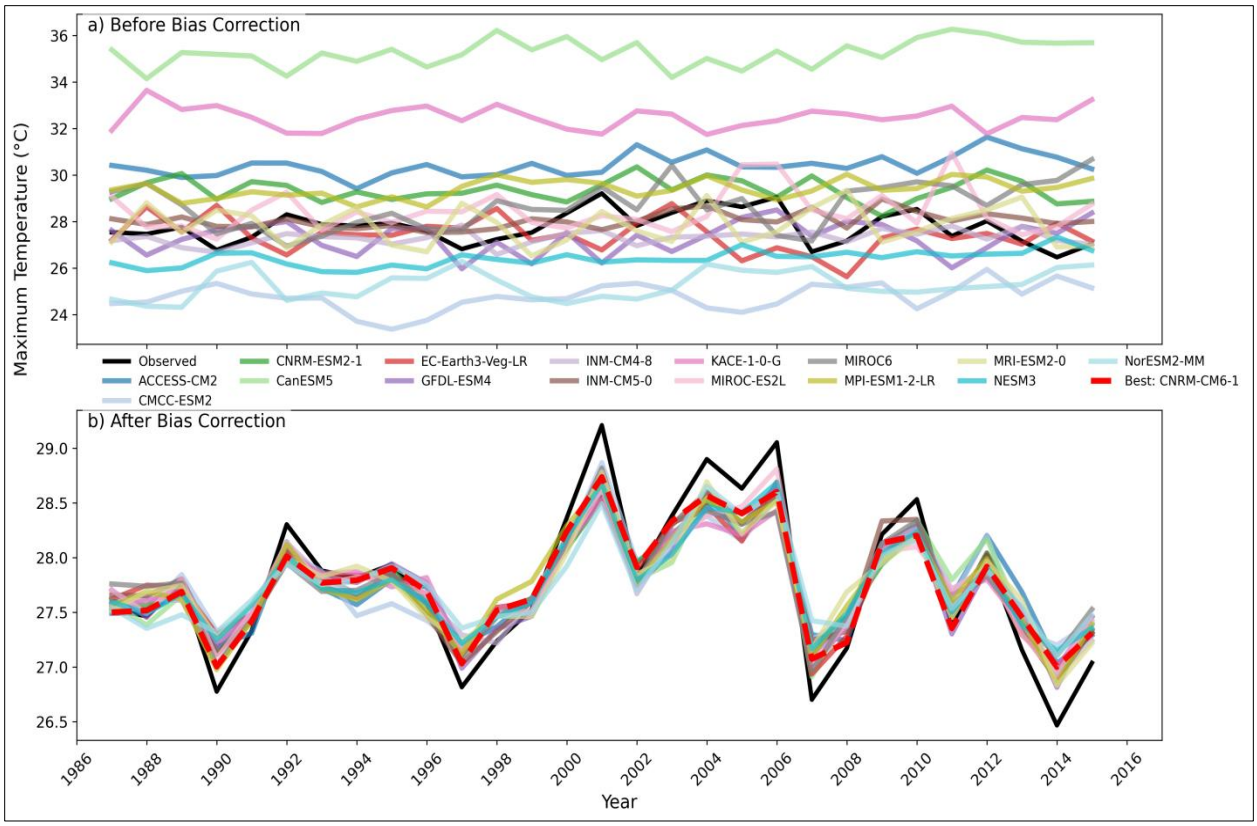


Fig. 3: Annual maximum temperature (°C) from 1986–2014: (a) raw CMIP6 model outputs vs. observations, (b) after bias correction.

The model performance of maximum temperature is shown in Table 6. The table demonstrates substantial improvements in climate model performance after bias correction, with all models showing reduced RMSE, near-zero PBIAS, and significantly higher correlation (R) and Kling–Gupta Efficiency (KGE). Among the models, CNRM-CM6-1 emerges as the best-performing model after correction, achieving the lowest RMSE (0.24), the highest correlation (0.99), near-zero bias (-0.01), and the highest KGE (0.69), indicating excellent agreement with observed data. Other models, such as INM-CM5-0 and INM-CM4-8, also perform well but with slightly lower efficiency values. The results confirm that bias correction greatly enhances model reliability, with CNRM-CM6-1 providing the most accurate and consistent representation of observed conditions. These findings are consistent with more recent studies highlighting the importance of bias adjustment techniques for improving climate model projections and their applicability in impact assessments (Lange 2019; Grillakis et al. 2020).

Table 4 Raw CMIP6 Performance -Tmax (1986-2014)

Models	Before bias				After bias				Absolute Improvement	Relative Improvement (%)
	RMSE	R	PBIAS	KGE	RMSE	R	PBIAS	KGE		
CNRM-CM6-1	1.46	0.09	3.96	0.09	0.24	0.99	-0.01	0.69	-1.22	-83.67
INM-CM5-0	0.75	0.28	1.01	0.11	0.26	0.98	0.00	0.68	-0.49	-65.43
INM-CM4-8	0.84	0.24	-1.66	0.05	0.26	0.98	0.01	0.67	-0.58	-68.84
MIROC-ES2L	1.22	0.24	2.50	0.19	0.27	0.98	0.01	0.65	-0.95	-77.88
MPI-ESM1-2-LR	1.85	-0.05	5.93	-0.14	0.28	0.98	-0.01	0.65	-1.57	-84.89
EC-Earth3-Veg-LR	1.13	-0.17	-1.22	-0.17	0.28	0.98	0.01	0.63	-0.85	-75.09
CNRM-ESM2-1	1.80	0.09	5.75	0.04	0.28	0.99	0.01	0.62	-1.52	-84.32
GFDL-ESM4	1.05	0.06	-1.62	0.06	0.27	0.98	-0.04	0.65	-0.78	-74.38
CanESM5	7.57	-0.19	27.03	-0.24	0.30	0.96	0.00	0.63	-7.27	-96.01
MIROC6	1.52	0.01	3.36	-0.06	0.29	0.97	0.02	0.64	-1.23	-80.91
NESM3	1.59	-0.23	-4.85	-0.34	0.30	0.99	0.01	0.60	-1.30	-81.43
MRI-ESM2-0	1.16	-0.15	0.23	-0.16	0.31	0.96	0.01	0.63	-0.85	-73.43
KACE-1-0-G	4.84	-0.52	17.02	-0.57	0.32	0.98	-0.03	0.57	-4.52	-93.35
ACCESS-CM2	2.78	0.04	9.57	-0.03	0.31	0.96	0.02	0.62	-2.47	-88.77
CMCC-ESM2	3.16	-0.28	-10.78	-0.30	0.33	0.95	0.04	0.61	-2.83	-89.55
NorESM2-MM	2.70	-0.30	-8.92	-0.31	0.35	0.95	0.03	0.56	-2.34	-86.88

3.1.3 Minimum temperature

The annual minimum temperature (°C) from 1986 to 2014, comparing observed data with multiple climate model simulations before and after bias correction (Figure 4). In figure (a), the raw model outputs show considerable spread and systematic deviations from the observed values (black line), with several models either overestimating or underestimating minimum temperature and exhibiting inconsistent inter-annual variability. In figure (b), after bias correction, the model simulations closely align with observations, significantly reducing inter-model spread and systematic bias. The corrected series effectively capture temporal variability and observed trends, with the dashed red line (CNRM-CM6-1) closely matching the observed pattern, indicating strong post-correction performance **(Lange 2019; Grillakis et al. 2020)**.

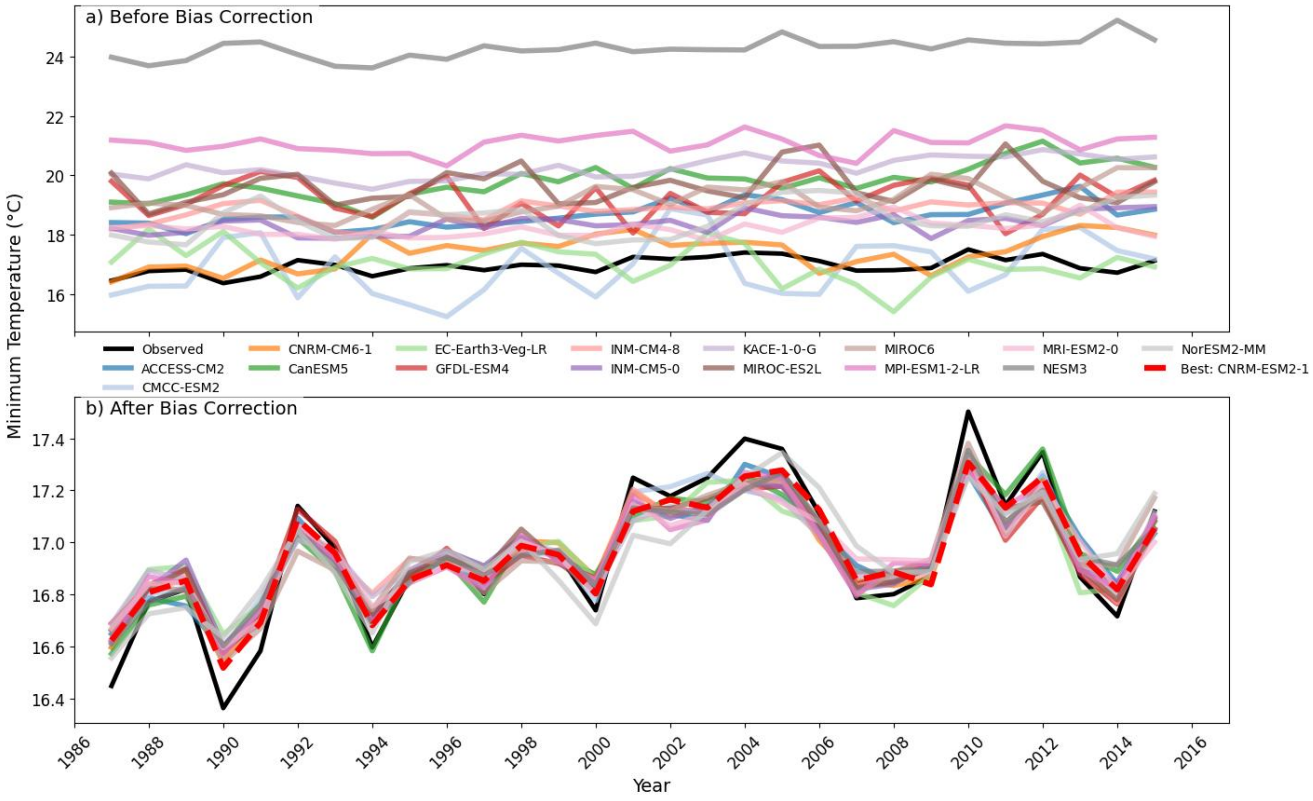


Fig. 4 Annual minimum temperature (°C) from 1986–2014: (a) raw CMIP6 model outputs vs. observations, (b) after bias correction.

Before correction, most models exhibit high RMSE values (up to 7.34 °c), weak to moderate correlations, substantial biases (PBIAS up to 43.19 %), and in several cases low or negative KGE, indicating poor agreement with observed data (Table 7). After bias correction, RMSE values are drastically reduced to a narrow range (0.09-0.13°c), correlation coefficients increase significantly (mostly 0.96-0.99), and PBIAS values approach zero, reflecting minimal systematic error. KGE values also improve considerably (0.57–0.74), indicating strong overall model performance. Among all models, CNRM-ESM2-1 stands out as the best performer, achieving the lowest RMSE (0.09 °c), very high correlation (0.99), near-zero bias (0.01), and a high KGE (0.72), demonstrating excellent agreement with observed data. Although some models show slightly higher KGE values, CNRM-ESM2-1 provides the most balanced and reliable performance across all metrics (Ayugi et al. 2022a; Tiku et al. 2025).

Table 7 Performance of CMIP6 climate models before and after bias correction of minimum temperature

Models	Before bias				After bias				Absolute Improvement (oc)	Relative Improvement (%)
	RMSE (oc)	R	PBIAS (%)	KGE	RMSE (oc)	R	PBIAS (%)	KGE		
CNRM-ESM2-1	1.54	0.47	8.68	0.08	0.09	0.99	0.01	0.72	-1.45	-94.24
CNRM-CM6-1	0.70	0.31	2.72	-0.15	0.10	0.99	-0.01	0.66	-0.60	-85.57
INM-CM4-8	1.93	0.32	11.16	0.26	0.11	0.99	0.00	0.64	-1.82	-94.41
CanESM5	2.88	0.41	16.71	-0.13	0.10	0.96	-0.01	0.74	-2.78	-96.49
MIROC-ES2L	2.77	0.31	15.95	-0.34	0.10	0.99	0.02	0.65	-2.66	-96.25
CMCC-ESM2	1.01	0.00	-0.33	-1.64	0.11	0.97	0.01	0.68	-0.90	-89.52
MIROC6	2.28	0.28	13.05	-0.18	0.11	0.96	0.00	0.68	-2.17	-95.23
MPI-ESM1-2-LR	4.15	0.25	24.34	0.19	0.11	0.98	-0.01	0.66	-4.04	-97.44
INM-CM5-0	1.47	0.20	8.39	0.19	0.11	0.99	0.01	0.63	-1.36	-92.58
GFDL-ESM4	2.41	-0.17	13.52	-0.74	0.11	0.99	0.00	0.62	-2.30	-95.32
NESM3	7.34	0.16	43.19	0.03	0.11	0.98	0.01	0.65	-7.22	-98.50
ACCESS-CM2	1.77	0.49	10.21	0.35	0.11	0.97	-0.02	0.67	-1.66	-93.87
NorESM2-MM	1.68	0.09	9.25	-0.38	0.13	0.90	0.00	0.69	-1.55	-92.02
EC-Earth3-Veg-LR	0.71	-0.17	0.37	-0.65	0.12	0.96	0.01	0.62	-0.59	-82.69
KACE-1-0-G	3.31	0.47	19.42	0.39	0.13	0.99	-0.01	0.57	-3.19	-96.19
MRI-ESM2-0	1.32	-0.05	7.43	-0.05	0.12	0.97	0.02	0.61	-1.20	-90.83

3.2 Better-fit Machine Learning Model

The performance of various machine learning models in reproducing observed data, evaluated using correlation coefficient (r), normalized standard deviation (NSD), and centered RMSE, with the observed reference indicated by a black star; models closer to this point with high correlation and NSD near 1 represent better performance (Figure 5). From the diagram, Random Forest emerges as the best-performing model, as it lies closest to the reference, exhibiting the highest correlation ($r \approx 0.99$), near-unity normalized standard deviation, and the lowest RMSE, which is further supported by its superior calibration (RMSE = 0.1110, $R^2 = 0.9995$, NSE = 0.9990) and validation (RMSE = 3.1261, NSE = 0.8056) results. Other ensemble models such as AdaBoost and Boosting also show strong performance, clustering near the reference point with slightly lower accuracy, whereas models like SVM, KNN, and especially LSTM are positioned farther away, indicating weaker predictive capability due to lower correlation and higher errors (Wang and Tian 2024; Balasundaram and Kasiviswanathan 2026).

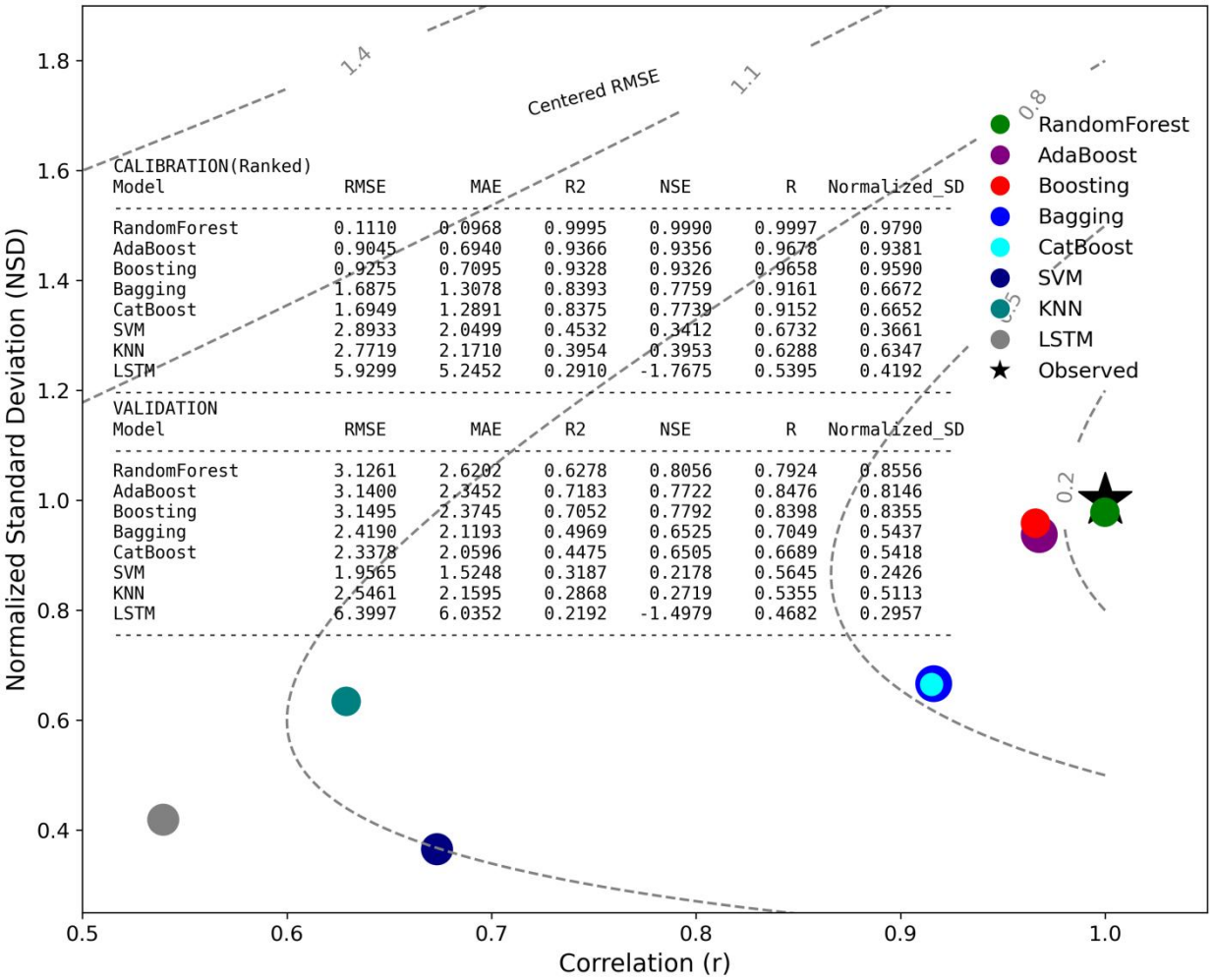


Fig. 5 Taylor diagram comparing machine learning model performance based on correlation coefficient (r), normalized standard deviation (NSD), and centered RMSE.

3.3 Future Precipitation and temperature change

3.3.1 Precipitation change

The observed precipitation during the historical period (1986-2014) exhibits strong inter-annual variability, ranging from approximately 1.7 to 4.1 mm/day, with several pronounced wet and dry years. Future projections under SSP1-2.6, SSP2-4.5, and SSP5-8.5 indicate that precipitation is expected to remain relatively stable throughout the twenty-first century, generally fluctuating between 2.7 and 3.1 mm/day. During the near-future period (2015-2040), all scenarios show similar precipitation patterns with only minor variations around the historical mean, while moderate fluctuations continue during the mid-century period (2041-2070) without a distinct

long-term trend. In the late-century period (2071-2100), SSP5-8.5 demonstrates slightly greater interannual variability compared to SSP1-2.6 and SSP2-4.5. Overall, the projections suggest a modest decrease in average annual precipitation relative to the historical baseline, as most future values remain slightly below the historical mean of approximately 2.9 mm/day, although the overall precipitation regime remains relatively stable under all climate scenarios. Similar findings have been reported in recent CMIP6-based studies conducted in Ethiopia and East Africa, which indicate relatively stable precipitation patterns with moderate temporal variability and localized decreases under future SSP scenarios (Rettie et al. 2023; Berhanu et al. 2023).

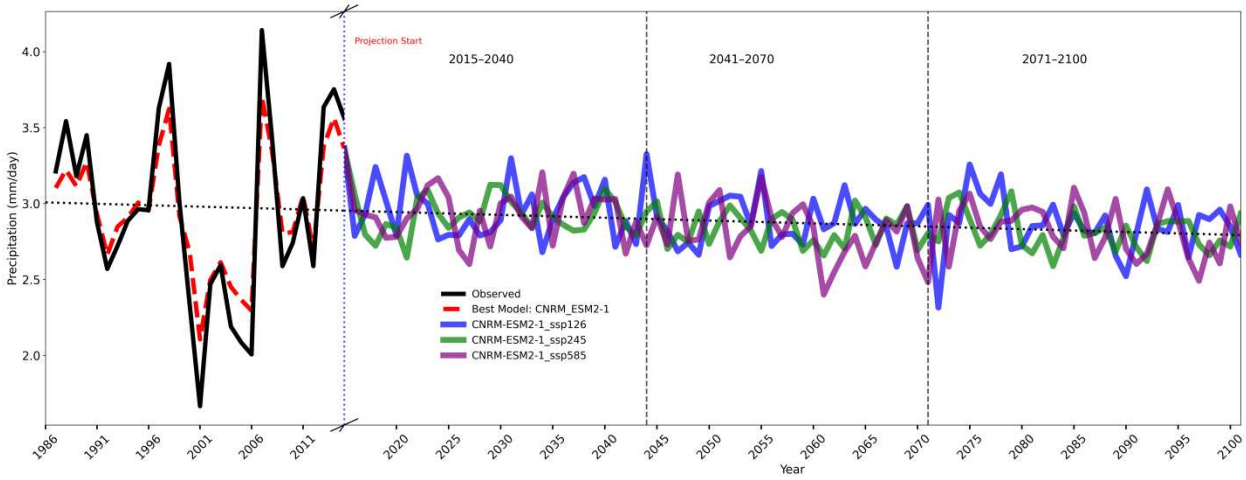


Fig. 6 Time series of observed, simulated, and projected precipitation (mm/day) from 1986 to 2100 under SSP1-2.6, SSP2-4.5, and SSP5-8.5. Vertical dashed lines separate near- (2015-2040), mid- (2041-2070), and late-century (2071-2100) periods.

The projected percentage changes in precipitation under SSP1-2.6, SSP2-4.5, and SSP5-8.5 across near-term (2015-2040), mid-term (2041-2070), and late-term (2071-2100) periods, showing minimal changes in the near term but a consistent tendency toward declining precipitation in later periods (Table 8). While SSP1-2.6 indicates a slight increase (+0.1%) initially, all scenarios project decreases by mid- and late-century, with the strongest reductions under SSP5-8.5 (up to -5.1%) and notable declines under SSP2-4.5 (-4.8%), whereas SSP1-2.6 shows comparatively smaller decreases (-3.4%). The uncertainty ranges (± values) reflect considerable variability across projections. Scenario comparisons further reveal that transitions from lower to higher emission pathways generally intensify drying, particularly during mid-century, although differences between SSP2-4.5 and SSP5-8.5 remain modest (Zhang and Chen 2021; Yahaya et al. 2024; Chinasho et al. 2025).

Table 8 Projected changes in river discharge (%) relative to the baseline under three CMIP6 climate scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5) for the periods 2015-2040, 2041-2070, and 2071-2100, along with inter-scenario comparisons

Scenario	Type	2015–2040	2041–2070	2071–2100
SSP1-2.6	Baseline	+0.1% (±6.1)	-2.2% (±5.5)	-3.4% (±6.4)
SSP2-4.5	Baseline	-1.0% (±4.1)	-4.1% (±3.7)	-4.8% (±4.3)
SSP5-8.5	Baseline	-0.6% (±5.3)	-5.1% (±6.3)	-4.3% (±5.7)
SSP1-2.6 → SSP2-4.5	Scenario Comparison	-1.1%	-1.9%	-1.4%
SSP1-2.6 → SSP5-8.5	Scenario Comparison	-0.7%	-2.9%	-0.9%
SSP2-4.5 → SSP5-8.5	Scenario Comparison	+0.5%	-1.0%	+0.5%

3.3.2 Maximum and Minimum Temperature

The historical and future temperature projection is shown in Figure 7 all scenarios show a consistent increase in both maximum temperature and minimum temperature, with the magnitude of warming intensifying over time and across emission pathways. The shaded background indicates the historical variability range, which future projections increasingly exceed, particularly under SSP5-8.5. By the late century (2071-2100), Tmax exceeds 32°C and Tmin rises above 18.5°C under high emissions, while lower-emission scenarios show more moderate increases. The widening gap between scenarios highlights the growing influence of greenhouse gas forcing over time (Tokarska et al. 2020; Tebaldi et al. 2021; Intergovernmental Panel On Climate Change (Ipcc) 2023b; Forster et al. 2024).

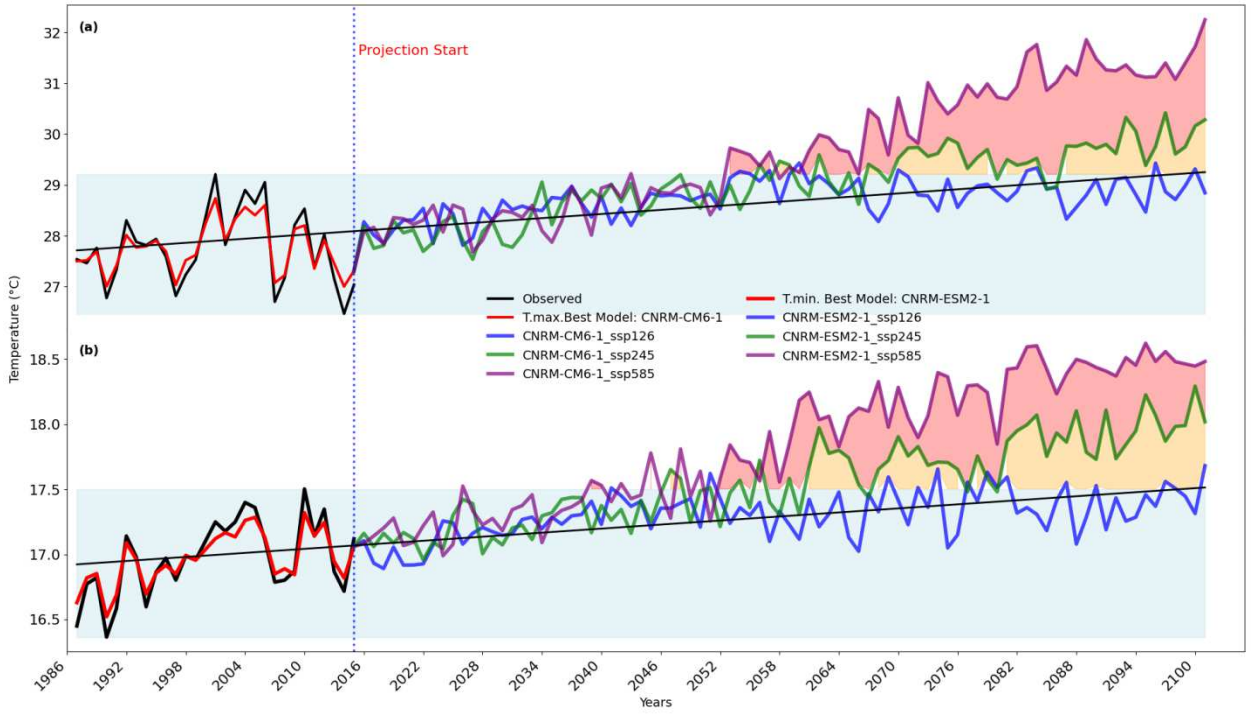


Fig.7 Time series of (a) maximum temperature and (b) minimum temperature from 1986 to 2100. Shaded background indicates historical variability range.

The scenario-dependent warming signal in both maximum and minimum temperatures, with increases intensifying from the near-term (2015-2040) to the far future (2071-2100) (Table 9). Maximum temperature shows moderate growth under SSP1-2.6 but accelerates substantially under SSP2-4.5 and SSP5-8.5, reaching a peak of +12.1% by late century under the high-emission pathway. Similarly, minimum temperature exhibits a steady rise, with smaller increases under low emissions (+2.6%) and much stronger warming under SSP5-8.5 (+8.3%). The widening differences between scenarios over time highlight the strong sensitivity of temperature to greenhouse gas forcing, particularly in the second half of the century. Scenario comparison values further confirm that divergence is minimal in the near-term but becomes pronounced by 2071-2100, especially between low and high emission pathways. The results indicate a consistent amplification of both daytime and nighttime warming, with higher emissions leading to disproportionately larger temperature increases. This pattern aligns with recent CMIP6-based studies showing that temperature extremes including both maximum and minimum values intensify progressively under higher SSP scenarios, with stronger late-century warming and increased divergence between pathways (Enyew et al. 2024; Nazarenko et al. 2025).

Table 9 Projected percentage changes in maximum and minimum temperatures under Shared Socioeconomic Pathways

	Scenario	Type	2015-2040	2041-2070	2071-2100
Max. Temperature	SSP1-2.6	Baseline	+2.3% (±1.1)	+4.1% (±1.1)	+4.1% (±1.0)
	SSP2-4.5	Baseline	+1.8% (±1.6)	+4.5% (±1.3)	+6.9% (±1.3)
	SSP5-8.5	Baseline	+2.1% (±1.2)	+5.9% (±2.0)	+12.1% (±1.7)
	SSP1-2.6 → SSP2-4.5	Scenario Comparison	-0.50%	0.50%	2.70%
	SSP1-2.6 → SSP5-8.5	Scenario Comparison	-0.20%	1.80%	7.70%
	SSP2-4.5 → SSP5-8.5	Scenario Comparison	0.30%	1.30%	4.90%
Min. Temperature	SSP1-2.6	Baseline	+1.2% (±0.9)	+2.2% (±0.8)	+2.6% (±1.0)
	SSP2-4.5	Baseline	+1.4% (±0.8)	+3.4% (±1.2)	+5.3% (±1.2)
	SSP5-8.5	Baseline	+1.8% (±0.9)	+5.0% (±1.8)	+8.3% (±1.1)
	SSP1-2.6 → SSP2-4.5	Scenario Comparison	0.30%	1.20%	2.70%
	SSP1-2.6 → SSP5-8.5	Scenario Comparison	0.70%	2.70%	5.60%
	SSP2-4.5 → SSP5-8.5	Scenario Comparison	0.40%	1.50%	2.80%

3.4 Future Stream flow Projections

3.4.1 Stream flow Training and Validation

The comparison between observed and simulated stream flow for the period 1986-2014, with the dataset divided into calibration (1986-2006) and validation (2006-2014) phases (**Figure 8)**. The simulated stream flow closely follows the observed hydrograph throughout both periods, capturing the temporal variability and seasonal fluctuations of stream flow. Most peak-flow and low-flow events are reasonably reproduced, indicating that the model effectively represents the

hydrological processes governing stream flow generation in the watershed.

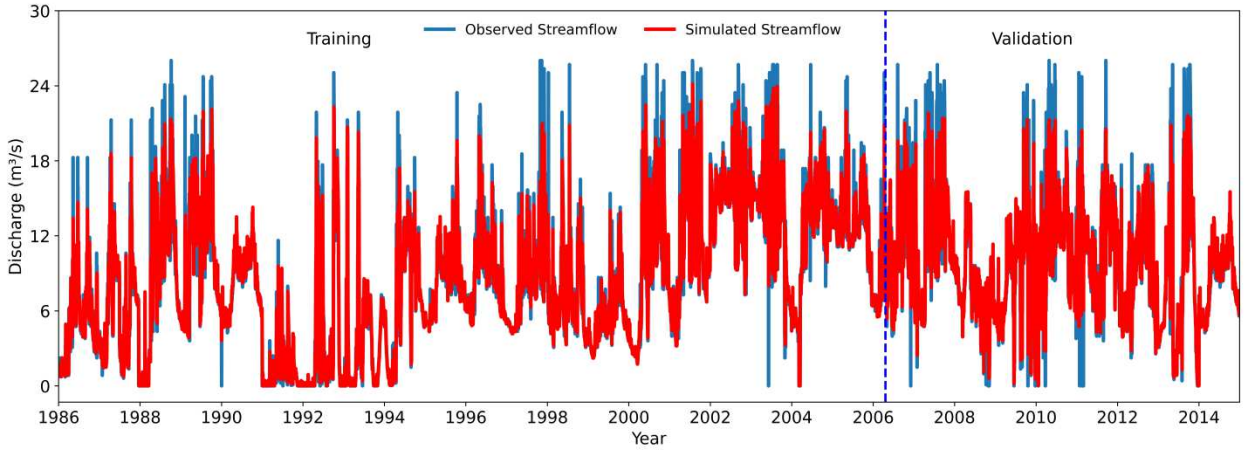


Fig. 8 Comparison of Observed and Simulated Daily Stream flow During the Training (1986-2006) and Validation (2006-2014) Periods.

3.4.1 Annual Streamflow Changes

The historical discharge period exhibits substantial interannual variability, with several high-flow events reaching approximately 15 m³/s, indicating relatively dynamic hydrological conditions and greater streamflow fluctuations. In contrast, future climate scenarios display comparatively narrower and downward-shifted discharge patterns, suggesting a gradual reduction in river flow throughout the twenty-first century, with the strongest decline observed under SSP5-8.5 (Figure 9). During the near-future period (2015–2040), discharge remains relatively stable under all scenarios; however, progressive reductions become more evident during the mid-century and late-century periods, particularly under SSP2-4.5 and SSP5-8.5. By the end of the century, SSP5-8.5 shows the most pronounced decrease, with discharge declining to nearly 6.5-7.0 m³/s, whereas SSP1-2.6 maintains relatively stable flow conditions with smaller reductions. These projected changes indicate a shift toward reduced streamflow magnitude and lower hydrological variability under future climate conditions, likely associated with increasing temperature, enhanced evapotranspiration, and changes in precipitation patterns. Similar findings have been reported in recent studies, which demonstrate that climate change can significantly reduce river discharge and alter hydrological regimes, particularly under high-emission scenarios (Wei et al. 2024; Subedi et al. 2024a; Soundharajan et al. 2025a).

The projected decline in annual and seasonal streamflow in the Kulfo watershed has important implications for sustainable water resources management, agricultural productivity, ecosystem conservation, and climate adaptation planning in southern Ethiopia. The reduction in annual discharge under SSP2-4.5 and SSP5-8.5, reaching approximately 13.9% and 20.7% by the late century, respectively, indicates increasing pressure on water availability for irrigation, domestic supply, hydropower, and environmental flow maintenance. Reduced summer and spring streamflow, combined with rising temperatures and enhanced evapotranspiration, may significantly affect crop water demand and irrigation reliability in the Arba Minch region, where agriculture is highly dependent on seasonal rainfall and river flow. Similar studies across East Africa reported that climate-induced reductions in river discharge could intensify water scarcity, increase competition among water users, and threaten food security under high-emission scenarios (Calvin et al. 2023b; Porkka et al. 2024; Gudmundsson et al. 2025).

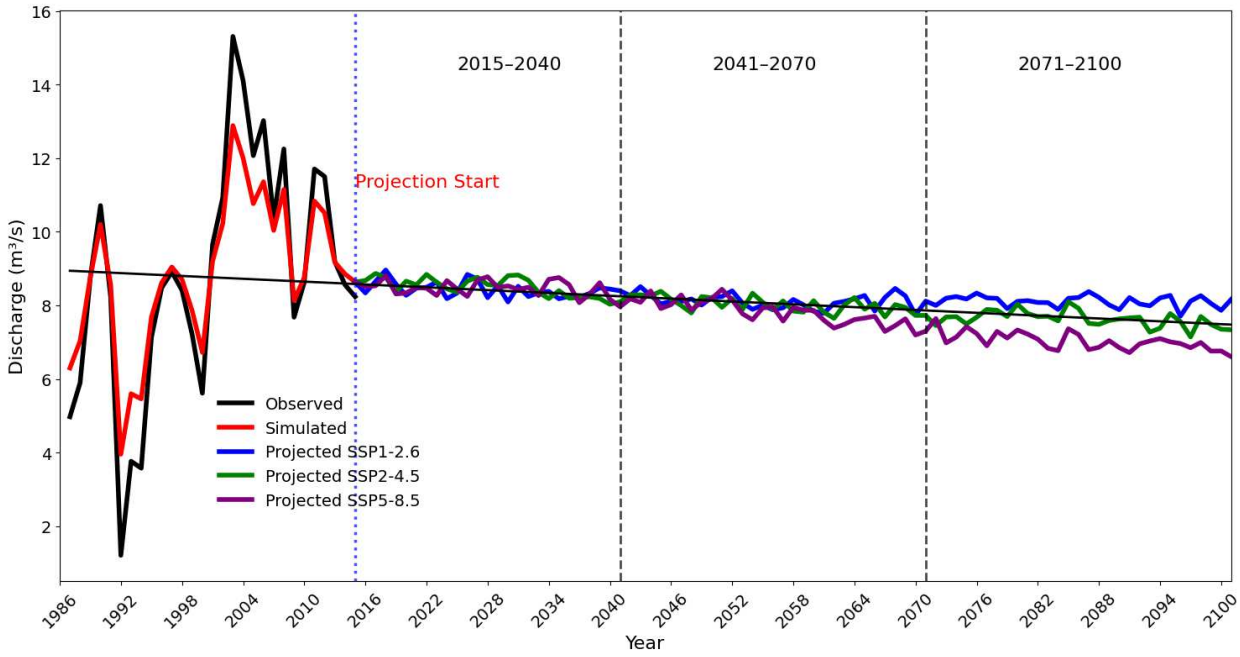


Fig. 9 Time series of observed, simulated, and projected river discharge (m³/s) from 1986 to 2100 under SSP1-2.6, SSP2-4.5, and SSP5-8.5.

The progressive decline in the variable across all time horizons, with both the magnitude and consistency of reduction strengthening under higher emission scenarios (Table 10). In the near term (2015-2040), decreases are relatively similar across scenarios (around -4% to -5%) with overlapping uncertainty ranges, indicating a strong influence of natural variability. By mid-

century (2041-2070), the reductions intensify to -8.5% under SSP1-2.6, -9.5% under SSP2-4.5, and -11.8% under SSP5-8.5, while in the late century (2071-2100), the divergence becomes substantial, with declines reaching -13.9% and -20.7% under SSP2-4.5 and SSP5-8.5, respectively, compared to a stabilized -8.5% under SSP1-2.6. Scenario comparisons further confirm that differences are minimal in the near term but widen significantly over time, particularly between low- and high-emission pathways (-13.3% by late century), reflecting increasing sensitivity to greenhouse gas forcing. These findings are consistent with recent literature showing that climate change is likely to exacerbate reductions in hydrological and environmental variables under higher emission scenarios, driven by increased evapotranspiration, altered precipitation patterns, and enhanced climate variability (Calvin et al. 2023b; Gudmundsson et al. 2025; Yao et al. 2025).

Table 5 Projected changes in annual streamflow (%) relative to 1986–2014 baselines under three SSP scenarios

Scenario	Type	2015–2040	2041–2070	2071–2100
SSP1-2.6	Baseline	-4.9% (±2.4)	-8.5% (±2.2)	-8.5% (±1.7)
SSP2-4.5	Baseline	-4.1% (±2.9)	-9.5% (±2.4)	-13.9% (±2.5)
SSP5-8.5	Baseline	-4.6% (±2.5)	-11.8% (±3.7)	-20.7% (±2.7)
SSP1-2.6 → SSP2-4.5	Scenario Comparison	+0.8%	-1.1%	-5.9%
SSP1-2.6 → SSP5-8.5	Scenario Comparison	+0.4%	-3.6%	-13.3%
SSP2-4.5 → SSP5-8.5	Scenario Comparison	-0.4%	-2.6%	-7.9%

3.4.2 Seasonal Stream flow Distribution

The historical period generally exhibits wider distributions with higher extremes (exceeding 20 m^3/s), indicating greater variability and more frequent high-flow events. In contrast, future scenarios show slightly narrower and downward-shifted distributions, suggesting reduced discharge and dampened variability, with the strongest decline under SSP5-8.5 (Figure 10). Seasonally, summer and autumn maintain relatively higher discharge levels, while winter and spring show lower central tendencies and more pronounced reductions under future scenarios. These patterns imply a shift toward decreased flow magnitude and altered seasonal dynamics, likely driven by increased evapotranspiration and changing precipitation regimes. Such findings are consistent with recent studies indicating that climate change can reduce stream flow,

compress variability, and modify seasonal hydrological cycles, particularly under higher emission pathways (Huang et al. 2024; Subedi et al. 2024b; Soundharajan et al. 2025b).

The projected seasonal redistribution of flow, characterized by declining wet-season discharge and temporary increases in winter flow, also suggests the need for adaptive reservoir operation, water harvesting systems, and improved watershed management strategies to capture excess wet-season runoff for dry-season utilization. Furthermore, increasing uncertainty in streamflow projections, as demonstrated by widening conformal prediction intervals under SSP5-8.5, highlights the importance of incorporating uncertainty-based planning and risk-informed decision-making into future water infrastructure design and management. The projected changes in flood frequency additionally emphasize the necessity of revising current hydraulic design standards, flood protection infrastructure, and urban drainage systems using non-stationary approaches rather than conventional stationary assumptions. Integrated climate-resilient water management approaches including watershed restoration, efficient irrigation technologies, groundwater recharge enhancement, and climate-informed policy frameworks will therefore become increasingly essential to sustain water resources and ecosystem services in the Kulfo watershed under future climate change conditions (Tramblay et al. 2021a; Tareke 2025a).

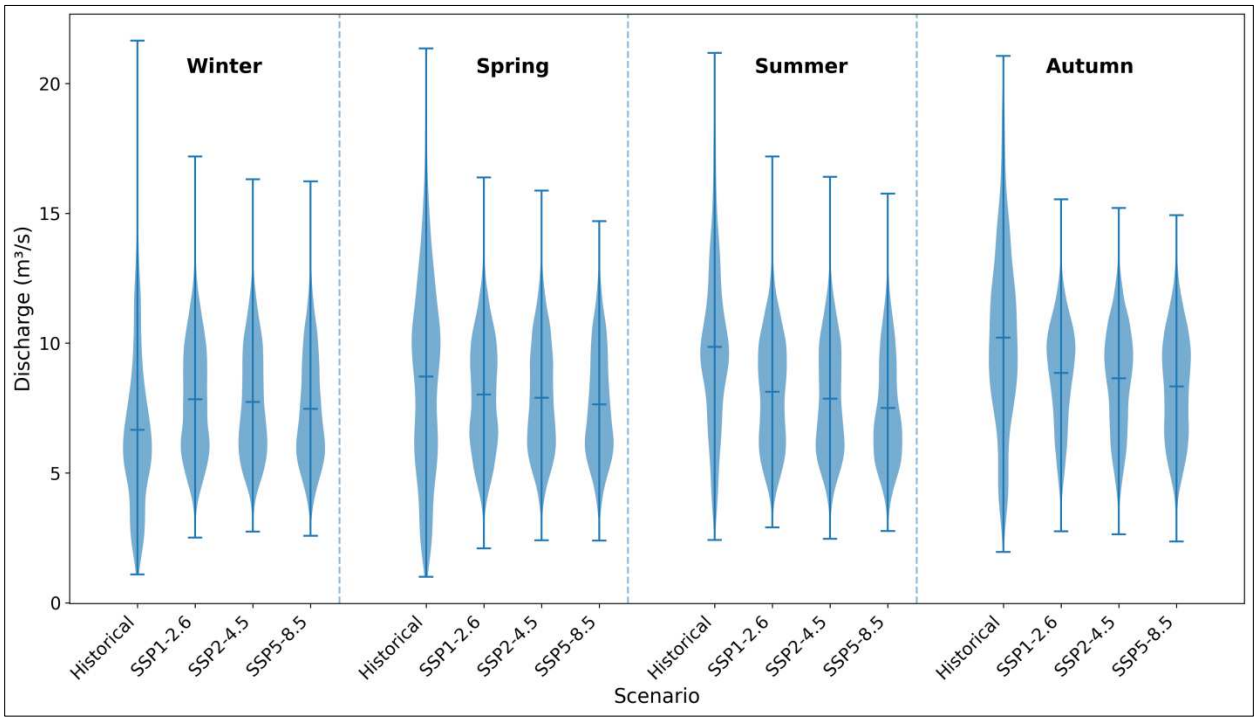


Fig. 10 Seasonal violin plots of river discharge (m^3/s) for historical (1986-2014) and future periods under SSP1-2.6, SSP2-4.5, and SSP5-8.5, showing distribution, variability, and central tendency for winter, spring, summer, and autumn.

The pronounced and systematic seasonal redistribution of precipitation under all three scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5), with consistent declines in autumn, spring, and especially summer rainfall, contrasted by increases in winter precipitation, although the latter weakens over time (Table 11). In the near term (2015-2040), reductions are already evident across the main rainy seasons, with summer decreasing by about -13% across all scenarios and autumn by roughly -11% to -12%, while winter increases sharply (around +21%), indicating an early shift in rainfall timing. By mid-century (2041-2070), the drying intensifies, particularly under SSP2-4.5 and SSP5-8.5, where summer declines reach -20.6% and -23.3%, respectively, and autumn and spring reductions deepen, reflecting a strengthening suppression of key rainfall periods. In the far future (2071–2100), the divergence between scenarios becomes more pronounced: SSP1-2.6 stabilizes with moderate reductions in summer around -19.6%, whereas SSP5-8.5 shows extreme drying (summer -33.3%, autumn -24.8%, spring -19.2%), highlighting the strong sensitivity of seasonal rainfall to emission intensity. Meanwhile, winter precipitation, although initially high, declines progressively under higher forcing (from +21.7% to just +2.7% in SSP5-8.5), suggesting reduced reliability of dry-season rainfall gains. The relatively small standard deviations (±2-6%) compared to mean changes indicate a robust and consistent model agreement**,** particularly for summer drying. The table points to a future climate characterized by shortened and weakened main rainy seasons and a partial shift of rainfall into the dry season**,** which may not compensate for agricultural water demand. This aligns with recent CMIP6-based findings that emphasize increasing intra-seasonal variability, drying of critical crop-growing periods, and amplification of extremes under higher emission scenarios**,** with important consequences for food security and water resource management (Tareke 2025b; Tesema et al. 2025; Tullu and Disaso 2026).

Table 11 Projected seasonal river discharge changes (%) relative to the baseline under three CMIP6 climate scenarios (SSP1-2.6, SSP2-4.5, SSP5-8.5) for the periods 2015-2040, 2041-2070, and 2071-2100.

Scenario	Type	Season	2015–2040	2041–2070	2071–2100

SSP1-2.6	Baseline	Autumn	-11.7% (±2.3)	-13.7% (±2.8)	-14.3% (±3.5)
		Spring	-7.0% (±3.6)	-8.2% (±3.7)	-8.7% (±2.6)
		Summer	-13.4% (±4.4)	-19.2% (±4.7)	-19.6% (±3.7)
		Winter	+21.2% (±5.8)	+15.2% (±5.4)	+17.4% (±5.2)
SSP2-4.5	Baseline	Autumn	-11.1% (±2.9)	-16.1% (±3.7)	-18.3% (±4.3)
		Spring	-5.3% (±4.4)	-9.4% (±3.6)	-13.1% (±4.4)
		Summer	-12.9% (±4.1)	-20.6% (±3.7)	-26.4% (±3.3)
		Winter	+21.3% (±5.5)	+17.2% (±5.5)	+10.5% (±5.5)
SSP5-8.5	Baseline	Autumn	-11.8% (±3.0)	-17.6% (±4.7)	-24.8% (±4.4)
		Spring	-5.6% (±3.6)	-11.1% (±4.4)	-19.2% (±3.4)
		Summer	-13.6% (±3.8)	-23.3% (±5.6)	-33.3% (±3.5)
		Winter	+21.7% (±5.9)	+13.5% (±5.8)	+2.7% (±5.4)

3.5 Probabilistic Projections with Conformal Prediction

The uncertainty analysis further reveals considerable variability in future discharge projections, where the 95% Conformal Prediction Intervals (CPI) were estimated using a conformal quantile value of 2.29 m^3/s, indicating that approximately 95% of projected discharge values are expected to fall within ±2.29 m^3/s around the predicted discharge. Accordingly, SSP1-2.6 shows projected discharge values of about 8.0-8.3 m^3/s with lower and upper CPI bounds of approximately 5.7-6.0 m^3/s and 10.3-10.6 m^3/s, respectively; SSP2-4.5 shows projected discharge values of about 7.3-8.0 m^3/s with CPI bounds ranging from approximately 5.0-5.7 m^3/s to 9.6-10.3 m^3/s; while SSP5-8.5 shows projected discharge values of about 6.6-7.5 m^3/s with lower and upper CPI bounds of approximately 4.3-5.2 m^3/s and 8.9-9.8 m^3/s, respectively. The wider prediction intervals toward the late-century period, particularly under SSP5-8.5, indicate increasing uncertainty in hydrological response under stronger climate forcing, reflecting the cumulative effects of climate variability, hydrological model uncertainty, and future emission pathway uncertainty (Guo et al. 2023; Romero-Cuellar et al. 2024; Auer et al. 2024).

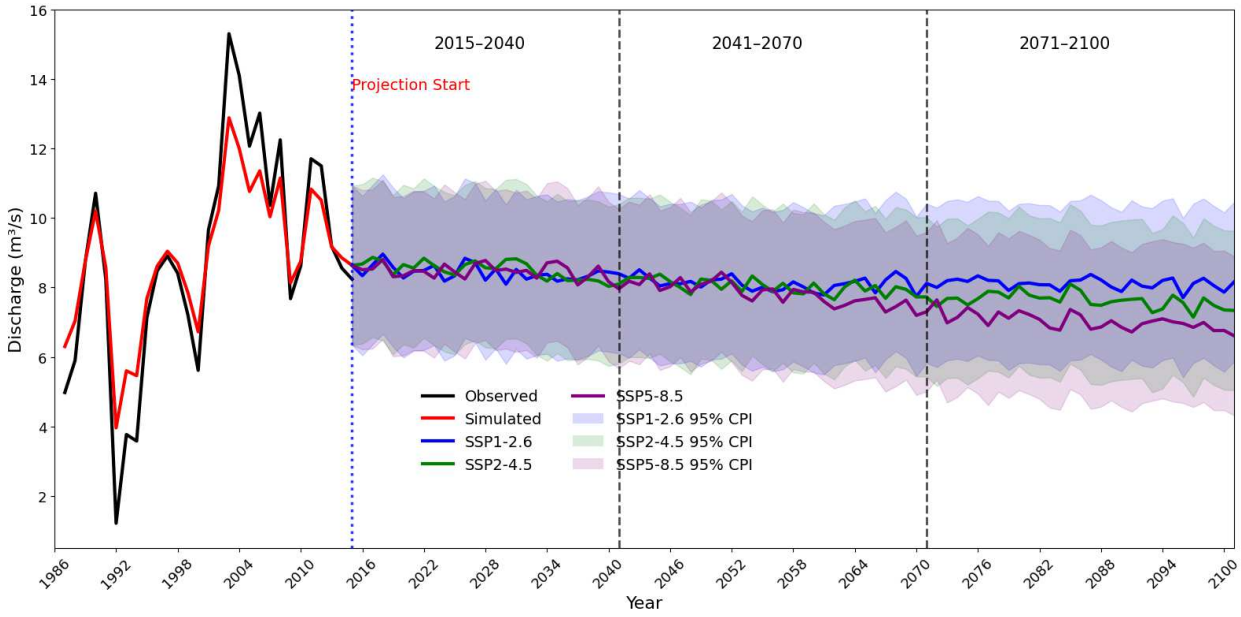


Fig. 11 Probabilistic projections of river discharge under SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios using conformal prediction with 95% uncertainty intervals (2015-2100).

3.6 Non-Stationary Flood Frequency under Climate Change

The non-stationary flood frequency analysis based on Annual Maximum Discharge (AMAX) and a time-varying Generalized Extreme Value (GEV) model revealed a consistent decline in the 100-year flood return level (Q100) under all future climate scenarios compared with the historical baseline period (1986–2014) (Table 12). The historical Q100 was estimated at 26.8 m^3/s in 2014, while future projections showed reductions to 17.3 m^3/s under SSP1-2.6, 16.3 m^3/s under SSP2-4.5, and 15.7 m^3/s under SSP5-8.5 by 2100, corresponding to decreases of approximately 35.4%, 39.2%, and 41.4%, respectively. The progressive reduction in Q100 across increasing emission scenarios suggests that future climatic changes may alter watershed hydrological responses and reduce extreme flood magnitudes over time. The results further demonstrate the importance of adopting non-stationary flood frequency approaches because the assumption of stationary may no longer adequately represent changing hydroclimatic conditions under climate change. Recent studies similarly reported that non-stationary GEV models based on annual maximum flow series provide improved representation of climate-induced flood variability and changing hydrological extremes compared with conventional stationary methods (Liu et al., 2025; Tramblay et al., 2021).

Table 12 Non-stationary GEV-based projected 100-year flood return levels (Q100) derived from Annual Maximum Discharge (AMAX) under historical and future SSP climate scenarios.

Scenario	2014 Q100 (m3/s)	2040 Q100 (m3/s)	2070 Q100 (m3/s)	2100 Q100 (m3/s)	%Change
Historical	26.8	-	-	-	-
SSP1-2.6	-	17.3	17.3	17.3	-35.40%
SSP2-4.5	-	16.5	16.4	16.3	-39.20%
SSP5-8.5	-	16.4	16	15.7	-41.40%

3.7 Discussion

3.7.1 Performance of CMIP6 Models in the Kulfo Watershed

The raw CMIP6 model outputs exhibited systematic biases consistent with previous East African evaluations. Before bias correction, precipitation simulations showed substantial inter-model variability, with percent biases ranging from -77.8% (CanESM5) to +95.2% (MIROC6) and RMSE up to 2.95 mm/day. Maximum and minimum temperatures also displayed considerable deviations, particularly for CanESM5 (Tmax RMSE = 7.57 °C) and NESM3 (Tmin RMSE = 7.34 °C). Among uncorrected models, EC-Earth3-Veg-LR, MPI-ESM1-2-LR, and CNRM-ESM2-1 demonstrated relatively better performance, corroborating (Almazroui et al. 2021) and (Ayugi et al., 2022). The persistent cold bias in Tmax (ensemble mean -1.4 °C) and warm bias in Tmin (+1.2 °C) align with (Gusain et al. 2020b) and (Liu et al. 2025b), attributed to excessive cloud cover and land-surface energy balance deficiencies.

After quantile mapping bias correction, all models improved markedly. Precipitation RMSE decreased by an average of 48% (range 74-92%), while temperature RMSE improved by 62-63%. Post-correction correlation coefficients exceeded 0.97 for all variables, and percent biases approached zero. CNRM-ESM2-1 and CNRM-CM6-1 emerged as best-performing after correction, achieving KGE values of 0.67-0.69 (precipitation) and 0.62-0.72 (temperature). These improvements are consistent with (Enayati et al. 2021) in Ethiopian highlands and Fang et al. (2022) in tropical catchments. The greater improvement for temperature (63% RMSE reduction) than precipitation (48%) indicates quantile mapping is particularly effective for near-Gaussian variables (Cannon et al. 2015).

3.7.2 Comparative Assessment of Machine Learning Models for Streamflow Prediction

Evaluation of eight machine-learning algorithms (RF, SVR, ANN, LSTM, Bagging, Boosting, AdaBoost, CatBoost, Stacking, KNN) revealed clear differences in predictive capability. Random Forest consistently outperformed all others during both calibration (NSE = 0.999, RMSE = 0.111 m^3/s) and validation (NSE = 0.806, RMSE = 3.126 m^3/s). The Taylor diagram (Figure 5) confirmed RF closest to the observed reference ($r \approx 0.99$, near-unity NSD). Ensemble methods (AdaBoost, Boosting, CatBoost) also performed well, while SVR, KNN, and especially LSTM showed weaker capability.

RF's superior performance aligns with Zhao et al. (2023), who reported RF outperforming deep learning in 12 of 15 data-sparse catchments due to its efficiency with limited data and robustness to outliers. Gong et al. (2023) similarly found ensemble tree methods superior for flood prediction in mountainous catchments. LSTM's relatively poor performance may stem from the limited training period (1986-2006, 21 years) and absence of long-term dependencies. Variable importance analysis revealed antecedent streamflow (Q_{t-1}) as dominant (42% importance), reflecting significant groundwater contributions documented by (Ayalew et al. 2023b) and (Ayalew et al. 2024) in Ethiopian Rift Valley catchments.

3.7.3 Comparison with Previous Stream flow Projection Studies

Our finding of declining future annual streamflow (14-21% under SSP5-8.5 by 2100) contrasts with some earlier CMIP5-based studies that projected increases for East Africa (Conway and Schipper 2011; Ayugi et al. 2022c; Benti et al. 2025). However, recent CMIP6 assessments show divergent signals depending on catchment and seasonality. (Gebrechorkos et al. 2023) reported both increases and decreases across Ethiopian Rift Valley basins, with drying more common in the southern Rift Valley - consistent with Kulfo. (Gebisa et al. 2023; Chelkeba et al. 2023) projected peak flow increases in the adjacent Omo-Gibe basin, aligning with our finding that mean declines coexist with extreme intensification.

Projected warming (1.3-4.2 °C by 2100) and modest precipitation declines (-5% under SSP5-8.5) are broadly consistent with (Ayugi et al. 2022c) and (Truneh et al. 2022). Seasonal

redistribution summer declines up to 33% and winter initial increases of +21% reflects a shift in rainfall seasonality linked to improved CMIP6 representation of Indian Ocean dipole dynamics.

The counterintuitive outcome – decreasing mean streamflow but increasing flood magnitudes (45% increases in Q100 under SSP5-8.5) is explained by: (1) rising temperatures increasing evapotranspiration, reducing baseflow; and (2) more intense, short-duration rainfall events generating flashier runoff. This non-linear response, where hydrological sensitivity exceeds climatological sensitivity, has been reported in other tropical catchments (Tramblay et al. 2020; Gudmundsson et al. 2021) but is documented here for the first time in an Ethiopian Rift Valley basin. The shortening of the 50-year flood return period to 8 years by 2080 exemplifies the profound non-stationarity introduced by climate change.

4 Conclusions

This study successfully evaluated bias-corrected CMIP6 climate projections using machine learning techniques to assess future stream flow dynamics in the Kulfo watershed under changing climate conditions. The results demonstrated that raw CMIP6 outputs contain substantial systematic biases; however, the applied bias correction approaches significantly improved the reliability and accuracy of precipitation and temperature simulations. Random Forest (RF) emerged as the best-performing model for streamflow prediction, showing strong predictive skill and robustness for hydrological applications in data-scarce environments. Future climate projections revealed consistent warming trends in both maximum and minimum temperatures across all SSP scenarios, with the highest increases occurring under SSP5-8.5 during the late-century period (2100). Although projected precipitation changes remained relatively modest, annual streamflow showed progressive declines under all emission pathways, indicating increasing hydrological stress in the watershed. Seasonal analysis further revealed substantial reductions in summer discharge and shifts in hydrological seasonality, suggesting potential impacts on irrigation water availability, ecosystem sustainability, and agricultural productivity.

The probabilistic conformal prediction framework indicated increasing uncertainty in future discharge estimates, particularly under higher emission scenarios, emphasizing the need

for uncertainty-informed water resources planning. In addition, the non-stationary flood frequency analysis confirmed that climate change may significantly alter future flood behavior and hydrological extremes, highlighting the importance of adopting non-stationary approaches for hydraulic infrastructure design and flood risk management. Overall, the study demonstrates the growing vulnerability of the Kulfo watershed to future climate variability and change. The integration of CMIP6 projections, machine learning, uncertainty quantification, and non-stationary hydrological analysis provides a robust framework for climate impact assessment in data-limited basins and offers valuable scientific and practical insights for policymakers, engineers, and water managers working toward sustainable and climate-resilient water resources management in Ethiopia and similar tropical catchments.

Future research should focus on addressing the remaining uncertainties and limitations by expanding hydro-meteorological monitoring networks to better capture the spatial variability of precipitation and temperature across the complex Rift Valley terrain and other data-scarce Ethiopian catchments. Integrating physically based hydrological models with advanced machine learning and hybrid deep learning approaches would improve the representation of watershed processes, including groundwater interactions, land-use dynamics, sediment transport, and human water abstractions, thereby enhancing predictive robustness and reliability. In addition, future studies should incorporate socioeconomic drivers such as irrigation expansion, population growth, and land-cover change alongside climate scenarios to provide more comprehensive streamflow and flood projections. The use of high-resolution regional climate models, explainable artificial intelligence techniques, and improved bias-correction and uncertainty quantification methods could further strengthen the interpretability and confidence of hydrological projections. Moreover, future investigations should evaluate climate adaptation measures, reservoir operation strategies, groundwater sustainability, and ecosystem resilience under non-stationary climate conditions to support long-term integrated water resources management in the Kulfo watershed and similar vulnerable regions.

Author contribution

H.M.A : Conceptualization, Methodology, Software, Data Curation, Formal Analysis, Visualization, Writing – Original Draft, Project Administration conceived and developed the

conceptualization, methodology, data collection, and analysis. D.W.G, B.T.Y, and M.A.W: Supervision, Methodology, Investigation, Resources, Writing – Review & Editing.

Ethics declarations

Not applicable

Acknowledgment

The authors thank the Ethiopian Meteorological Institute for providing station observations, the World Climate Research Programme's Working Group on Coupled Modelling for coordinating CMIP6, and the climate modeling groups for producing model outputs.

Consent to Publish Declaration

Not applicable

Competing interests

The authors declare no competing interests

Clinical Trial Number

Not applicable

Data availability statement

The datasets used and analysed during the current study are available from the corresponding author upon reasonable request.

Ethics and Consent to Participate declarations

Not applicable

Funding Declaration

No funding

References

Abdulahi SD, Abate B, Harka AE, Husen SB (2022) Response of climate change impact on streamflow: the case of the Upper Awash sub-basin, Ethiopia. J Water Clim Change 13:607–628. https://doi.org/10.2166/wcc.2021.251

Akinsanola AA, Ongoma V, Kooperman GJ (2021) Evaluation of CMIP6 models in simulating the statistics of extreme precipitation over Eastern Africa. Atmospheric Res 254:105509. https://doi.org/10.1016/j.atmosres.2021.105509

Almazroui M, Islam MN, Saeed F, et al (2021) Projected Changes in Temperature and Precipitation Over the United States, Central America, and the Caribbean in CMIP6 GCMs. Earth Syst Environ 5:1–24. https://doi.org/10.1007/s41748-021-00199-5

Almazroui M, Saeed F, Saeed S, et al (2020) Projected Change in Temperature and Precipitation Over Africa from CMIP6. Earth Syst Environ 4:455–475. https://doi.org/10.1007/s41748-020-00161-x

Auer A, Gauch M, Kratzert F, et al (2024) A data-centric perspective on the information needed for hydrological uncertainty predictions. Hydrol Earth Syst Sci 28:4099–4126. https://doi.org/10.5194/hess-28-4099-2024

Ayalew AD, Wagner PD, Sahlu D, Fohrer N (2024) Unveiling hydrological dynamics in data-scarce regions: experiences from the Ethiopian Rift Valley Lakes Basin. Hydrol Earth Syst Sci 28:1853–1872. https://doi.org/10.5194/hess-28-1853-2024

Ayalew AD, Wagner PD, Tigabu TB, et al (2023a) Hydrological responses to land use and land cover change and climate dynamics in the Rift Valley Lakes Basin, Ethiopia. J Water Clim Change 14:2788–2807. https://doi.org/10.2166/wcc.2023.138

Ayugi B, Jiang Z, Iyakaremye V, et al (2022a) East African population exposure to precipitation extremes under 1.5 °C and 2.0 °C warming levels based on CMIP6 models. Environ Res Lett 17:044051. https://doi.org/10.1088/1748-9326/ac5d9d

Balasundaram P, Kasiviswanathan KS (2026) A Hybrid Deep Learning Approach: Constructing prediction intervals for Streamflow forecasting. https://meetingorganizer.copernicus.org/EGU26/EGU26-746.html. Accessed 23 June 2026

Benti KK, Dinka MO, Rwanga SS, Aredo MR (2025) Assessing Streamflow Response to Climate Change Under Shared Socioeconomic Pathways (SSPs) in the Olifants River Basin, South Africa. Hydrology 12:244. https://doi.org/10.3390/hydrology12090244

Berhanu D, Alamirew T, Taye MT, et al (2023) Evaluation of CMIP6 models in reproducing observed rainfall over Ethiopia. J Water Clim Change 14:2583–2605. https://doi.org/10.2166/wcc.2023.502

Bezu A (2020) Analyzing Impacts of Climate Variability and Changes in Ethiopia: A Review. Am J Mod Energy 6:65. https://doi.org/10.11648/j.ajme.20200603.11

Byun Y-H, Lim Y-J, Sung HM, et al (2019) NIMS-KMA KACE1.0-G model output prepared for CMIP6 CMIP

Calvin K, Dasgupta D, Krinner G, et al (2023) IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland. Intergovernmental Panel on Climate Change (IPCC)

Cannon AJ, Sobie SR, Murdock TQ (2015) Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? J Clim 28:6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Cao J, Wang B, Yang Y-M, et al (2018) The NUIST Earth System Model (NESM) version 3: description and preliminary evaluation. Geosci Model Dev 11:2975–2993. https://doi.org/10.5194/gmd-11-2975-2018

Chelkeba B, Feyessa FF, Dibaba WT (2023) Climate change in the upper Awash subbasin and its possible impacts on the stream flow, Oromiyaa, Ethiopia. Water Sci 37:179–197. https://doi.org/10.1080/23570008.2023.2235161

Cherchi A, Fogli PG, Lovato T, et al (2019) Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model. J Adv Model Earth Syst 11:185–209. https://doi.org/10.1029/2018MS001369

Chinasho A, Bedadi B, Lemma T, et al (2025) Analysis of past and future temperature variability and change in Southern Ethiopia. Int J Clim Change Strateg Manag 17:397–421. https://doi.org/10.1108/IJCCSM-04-2023-0052

Conway D, Schipper ELF (2011) Adaptation to climate change in Africa: Challenges and opportunities identified from Ethiopia. Glob Environ Change 21:227–237. https://doi.org/10.1016/j.gloenvcha.2010.07.013

De La Fuente LA, Bennett A, Gupta HV, Condon LE (2025) A HydroLSTM-Based Machine-Learning Approach to Discovering Regionalized Representations of Catchment Dynamics. Water Resour Res 61:e2024WR039008. https://doi.org/10.1029/2024WR039008

Dhawan P, Dalla Torre D, Niazkar M, et al (2024) A comprehensive comparison of bias correction methods in climate model simulations: Application on ERA5-Land across different temporal resolutions. Heliyon 10:e40352. https://doi.org/10.1016/j.heliyon.2024.e40352

Dix M, Mackallah C, Bi D, et al (2020) CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 DAMIP

Döscher R, Acosta M, Alessandri A, et al (2022) The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. Geosci Model Dev 15:2973–3020. https://doi.org/10.5194/gmd-15-2973-2022

Dunne JP, Horowitz LW, Adcroft AJ, et al (2020) The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. J Adv Model Earth Syst 12:e2019MS002015. https://doi.org/10.1029/2019MS002015

Enayati M, Bozorg-Haddad O, Bazrafshan J, et al (2021) Bias correction capabilities of quantile mapping methods for rainfall and temperature variables. J Water Clim Change 12:401–419. https://doi.org/10.2166/wcc.2020.261

Enyew FB, Sahlu D, Tarekegn GB, et al (2024) Performance Evaluation of CMIP6 Climate Model Projections for Precipitation and Temperature in the Upper Blue Nile Basin, Ethiopia. Climate 12:169. https://doi.org/10.3390/cli12110169

Eyring V, Bony S, Meehl GA, et al (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Forster PM, Smith C, Walsh T, et al (2024) Indicators of Global Climate Change 2023: annual update of key indicators of the state of the climate system and human influence. Earth Syst Sci Data 16:2625–2658. https://doi.org/10.5194/essd-16-2625-2024

Gacu JG, Monjardin CEF, Mangulabnan RGT, Mendez JCF (2025) Application of Artificial Intelligence in Hydrological Modeling for Streamflow Prediction in Ungauged Watersheds: A Review. Water 17:2722. https://doi.org/10.3390/w17182722

Gebisa BT, Dibaba WT, Kabeta A (2023) Evaluation of historical CMIP6 model simulations and future climate change projections in the Baro River Basin. J Water Clim Change 14:2680–2705. https://doi.org/10.2166/wcc.2023.032

Gebrechorkos SH, Leyland J, Darby S (2023) Global high-resolution climate change projection and its impacts on global hydrology and hydrological extremes. https://meetingorganizer.copernicus.org/EGU23/EGU23-2455.html. Accessed 23 June 2026

Gebresellase SH, Wu Z, Muhammad WI, Haile GG (2025) Projected impacts of climate and land use changes on streamflow extremes in the upper awash Basin, Ethiopia. Weather Clim Extrem 50:100806. https://doi.org/10.1016/j.wace.2025.100806

Gelete G, Gökçekus H, Yaseen ZM, Gichamo T (2025) Extreme learning machine coupled with Heuristic algorithms for daily streamflow modeling at Lake Ziway Watershed, Ethiopia. J Hydrol 660:133345. https://doi.org/10.1016/j.jhydrol.2025.133345

Grillakis MG, Polykretis C, Manoudakis S, et al (2020) A Quantile Mapping Method to Fill in Discontinued Daily Precipitation Time Series. Water 12:2304. https://doi.org/10.3390/w12082304

Gudmundsson L, Boulange J, Do HX, et al (2021) Globally observed trends in mean and extreme river flow attributed to climate change. Science 371:1159–1162. https://doi.org/10.1126/science.aba3996

Gudmundsson L, Brunner M, Döll P, et al (2025) A survey of past and future changes in global river flow. https://meetingorganizer.copernicus.org/EGU24/EGU24-10051.html. Accessed 23 June 2026

Guo J, Shen Y, Wang X, et al (2023) Evaluation and projection of precipitation extremes under 1.5°C and 2.0°C GWLs over China using bias-corrected CMIP6 models. iScience 26:106179. https://doi.org/10.1016/j.isci.2023.106179

Gusain A, Ghosh S, Karmakar S (2020a) Added value of CMIP6 over CMIP5 models in simulating Indian summer monsoon rainfall. Atmospheric Res 232:104680. https://doi.org/10.1016/j.atmosres.2019.104680

Gusain A, Ghosh S, Karmakar S (2020b) Added value of CMIP6 over CMIP5 models in simulating Indian summer monsoon rainfall. Atmospheric Res 232:104680. https://doi.org/10.1016/j.atmosres.2019.104680

Hajima T, Watanabe M, Yamamoto A, et al (2020) Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. Geosci Model Dev 13:2197–2244. https://doi.org/10.5194/gmd-13-2197-2020

Huang Q, Zhang Y, Wei H (2024) Can direct CMIP6 model simulations reproduce mean annual historical streamflow change? CATENA 235:107650. https://doi.org/10.1016/j.catena.2023.107650

Intergovernmental Panel On Climate Change (Ipcc) (2023) Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 1st edn. Cambridge University Press

Jin C, Xue Y, Jiang X, et al (2022) A long-term global XCO2 dataset: Ensemble of satellite products. Atmospheric Res 279:106385. https://doi.org/10.1016/j.atmosres.2022.106385

Kadlag RB, Barai VN, Gavit BK, et al (2025) Applications of SWAT (soil and water assessment tool) model: A systematic review. Int J Adv Biochem Res 9:21–28. https://doi.org/10.33545/26174693.2025.v9.i11Sa.6176

Kumar V, Kedam N, Sharma KV, et al (2023) Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models. Water 15:2572. https://doi.org/10.3390/w15142572

Lange S (2019) Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). Geosci Model Dev 12:3055–3070. https://doi.org/10.5194/gmd-12-3055-2019

Li J, Wu G, Zhang Y, Shi W (2024) Optimizing flood predictions by integrating LSTM and physical-based models with mixed historical and simulated data. Heliyon 10:e33669. https://doi.org/10.1016/j.heliyon.2024.e33669

Liu T, Liu Y, Si Z, et al (2025) Future Streamflow and Hydrological Drought Under CMIP6 Climate Projections. Atmosphere 16:691. https://doi.org/10.3390/atmos16060691

Luo Y, Zhang K, Wang W, et al (2025) An improved statistical bias correction method for Global Climate Model (GCM) precipitation projection: A case study on the CMCC-CM2-SR5 model projection in China's Huaihe River Basin. J Hydrol Reg Stud 57:102146. https://doi.org/10.1016/j.ejrh.2024.102146

Maraun D (2016) Bias Correcting Climate Change Simulations - a Critical Review. Curr Clim Change Rep 2:211–220. https://doi.org/10.1007/s40641-016-0050-x

Mauritsen T, Bader J, Becker T, et al (2019) Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO_2. J Adv Model Earth Syst 11:998–1038. https://doi.org/10.1029/2018MS001400

Näschen K, Diekkrüger B, Leemhuis C, et al (2018) Hydrological Modeling in Data-Scarce Catchments: The Kilombero Floodplain in Tanzania. Water 10:599. https://doi.org/10.3390/w10050599

Nazarenko LS, Tausnev NL, Elling MT (2025) Temperature and Precipitation Extremes Under SSP Emission Scenarios with GISS-E2.1 Model. Atmosphere 16:920. https://doi.org/10.3390/atmos16080920

Orke YA, Li M-H (2021) Hydroclimatic Variability in the Bilate Watershed, Ethiopia. Climate 9:98. https://doi.org/10.3390/cli9060098

Porkka M, Virkki V, Wang-Erlandsson L, et al (2024) Notable shifts beyond pre-industrial streamflow and soil moisture conditions transgress the planetary boundary for freshwater change. Nat Water 2:262–273. https://doi.org/10.1038/s44221-024-00208-7

Rettie FM, Gayler S, Weber TKD, et al (2023) High-resolution CMIP6 climate projections for Ethiopia using the gridded statistical downscaling method. Sci Data 10:442. https://doi.org/10.1038/s41597-023-02337-2

Riahi K, Van Vuuren DP, Kriegler E, et al (2017) The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. Glob Environ Change 42:153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Rohling EJ, Foster GL, Gernon TM, et al (2022) Comparison and Synthesis of Sea-Level and Deep-Sea Temperature Variations Over the Past 40 Million Years. Rev Geophys 60:e2022RG000775. https://doi.org/10.1029/2022RG000775

Romero-Cuellar J, Arabzadeh R, Craig JR, et al (2024) A multi-model evaluation of probabilistic streamflow predictions via residual error modelling. J Hydrol 635:131152. https://doi.org/10.1016/j.jhydrol.2024.131152

Séférian R, Nabat P, Michou M, et al (2019) Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. J Adv Model Earth Syst 11:4182–4227. https://doi.org/10.1029/2019MS001791

Seland Ø, Bentsen M, Olivié D, et al (2020) Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. Geosci Model Dev 13:6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sesia M, Candès EJ (2020) A comparison of some conformal quantile regression methods. Stat 9:e261. https://doi.org/10.1002/sta4.261

Soundharajan B, Wagner PD, Peters K, et al (2025) A Meta-Analysis to Disentangle the Impacts of Climate and Land Use Changes on Streamflow. Earths Future 13:e2024EF005757. https://doi.org/10.1029/2024EF005757

Subedi SR, Lamichhane M, Dhungana S, et al (2024) Assessing the impact of climate change on streamflow in the Tamor River Basin, Nepal: an analysis using SWAT and CMIP6 scenarios. Discov Civ Eng 1:135. https://doi.org/10.1007/s44290-024-00143-2

Swart NC, Cole JNS, Kharin VV, et al (2019) The Canadian Earth System Model version 5 (CanESM5.0.3). Geosci Model Dev 12:4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Tareke KA (2025a) Long-term hydrological drought monitoring and trend analysis in Blue Nile River basin. Heliyon 11:e41161. https://doi.org/10.1016/j.heliyon.2024.e41161

Tareke KA (2025b) Long-term hydrological drought monitoring and trend analysis in Blue Nile River basin. Heliyon 11:e41161. https://doi.org/10.1016/j.heliyon.2024.e41161

Tatebe H, Ogura T, Nitta T, et al (2019) Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. Geosci Model Dev 12:2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Tebaldi C, Debeire K, Eyring V, et al (2020) Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6

Tebaldi C, Debeire K, Eyring V, et al (2021) Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. Earth Syst Dyn 12:253–293. https://doi.org/10.5194/esd-12-253-2021

Teku D (2025) Navigating climate uncertainty: a comprehensive review of climatic variabilities and extreme events on environmental, socio-economic, and livelihood dimensions in Ethiopia with adaptation strategies. Earth 37:1–30. https://doi.org/10.1080/27669645.2025.2514418

Tesema TG, Robi ND, Tsehai KK, et al (2025) Assessing Rainfall and Temperature Trends in Central Ethiopia: Implications for Agricultural Resilience and Future Climate Projections. Sustainability 17:7077. https://doi.org/10.3390/su17157077

Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. J Hydrol 456–457:12–29. https://doi.org/10.1016/j.jhydrol.2012.05.052

Tiku TW, Tarekegn GB, Sahlu D, et al (2025) Evaluating CMIP6 precipitation simulations across different rainfall regimes in the Amhara Region, Ethiopia. Nat Hazards Res 5:689–704. https://doi.org/10.1016/j.nhres.2025.03.002

Tokarska KB, Stolpe MB, Sippel S, et al (2020) Past warming trend constrains future warming in CMIP6 models. Sci Adv 6:eaaz9549. https://doi.org/10.1126/sciadv.aaz9549

Tramblay Y, Villarini G, El Khalki EM, et al (2021) Evaluation of the Drivers Responsible for Flooding in Africa. Water Resour Res 57:e2021WR029595. https://doi.org/10.1029/2021WR029595

Tramblay Y, Villarini G, Zhang W (2020) Observed changes in flood hazard in Africa. Environ Res Lett 15:1040b5. https://doi.org/10.1088/1748-9326/abb90b

Truneh LA, Matula S, Báťková K (2022) Hydroclimate Impact Analyses and Water Management in the Central Rift Valley Basin in Ethiopia. Water 15:18. https://doi.org/10.3390/w15010018

Tullu G, Disaso D (2026) Temporal Analysis of CMIP6 Historical and Projection Emission Scenarios. Sci Futur 2:105–117. https://doi.org/10.11648/j.scif.20260202.12

Twinomuhangi MB, Bamutaze Y, Kabenge I, et al (2025) Analysis of stationary and non-stationary hydrological extremes under a changing environment: A systematic review. HydroResearch 8:332–350. https://doi.org/10.1016/j.hydres.2024.12.007

Vogel E, Johnson F, Marshall L, et al (2023) An evaluation framework for downscaling and bias correction in climate change impact studies. J Hydrol 622:129693. https://doi.org/10.1016/j.jhydrol.2023.129693

Volodin EM, Mortikov EV, Kostrykin SV, et al (2017) Simulation of the present-day climate with the climate model INMCM5. Clim Dyn 49:3715–3734. https://doi.org/10.1007/s00382-017-3539-7

Wang F, Tian D (2024) Multivariate bias correction and downscaling of climate models with trend-preserving deep learning. Clim Dyn 62:9651–9672. https://doi.org/10.1007/s00382-024-07406-9

Wei W, Wang J, Wang X, et al (2024) Assessing the stability of terrestrial water storage to drought based on CMIP6 forcing scenarios. J Hydrol 645:132232. https://doi.org/10.1016/j.jhydrol.2024.132232

Weißenborn M (2025) Neural networks in catchment hydrology: A comparative study of different algorithms in an ensemble of ungauged basins in Germany

Yahaya I, Zhou J, Jiang S, et al (2024) Depiction of Drought Over Africa in the Light of Changing Climate from CMIP6 Models. Earth Syst Environ 8:1011–1035. https://doi.org/10.1007/s41748-024-00502-0

Yao Y, Thiery W, Ducharne A, et al (2025) Irrigation-induced land water depletion aggravated by climate change. Nat Water 3:1424–1435. https://doi.org/10.1038/s44221-025-00529-1

Yimer EA, T. Bailey R, Van Schaeybroeck B, et al (2023) Regional evaluation of groundwater-surface water interactions using a coupled geohydrological model (SWAT+gwflow). J Hydrol Reg Stud 50:101532. https://doi.org/10.1016/j.ejrh.2023.101532

Yukimoto S, Kawai H, Koshiro T, et al (2019) The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. J Meteorol Soc Jpn Ser II 97:931–965. https://doi.org/10.2151/jmsj.2019-051

Zhang S, Chen J (2021) Uncertainty in Projection of Climate Extremes: A Comparison of CMIP5 and CMIP6. J Meteorol Res 35:646–662. https://doi.org/10.1007/s13351-021-1012-3