Supplementary Part

1. Methodology

1.1 Preprocessing, Type, Sources, and Quality of Data

Data from five meteorological stations, namely Delomena(DLM), Kibremengist(KBRM), Negele(NGL), Bore(BOR), and Hagereselam(HGRS) (1985-2014), were obtained from the Ethiopian National Meteorological Institute (ENMI).). Temperature and precipitation data were recorded at each station, as indicated in Table S1, except for Bore. The Bore station temperature data were obtained from the Climate Hazards Center InfraRed Temperature with Stations (CHIRTS Moreover, to obtain good spatial coverage, gridded reanalyses of rainfall and temperature were extracted from CHRIPS and CHIRTS. This is a high-resolution observational product used to validate temperature and precipitation (Verdin et al., 2020).

Table S1. Station names and missing extents of climate parameters

S/No.	Station Name	Collected period	% of Missing Daily Recorded Climate Data of Major Climate Stations in GDRB for the selected station			
			PCP	Tmax	Tmin	Remark
1	Delomena(DLM)	1985 to 2014	17.50	27.64	31.21	Filled
2	Kibremengist(KBRM)	1985 to 2014	13	16.53	11.84	Filled
3	Negele(NGL)	1985 to 2014	5.0	35.36	35.37	Filled
4	Bore(BOR)	1985 to 2014	12.35	NA	NA	Filled
5	Hagere selam(HGRS)	1985 to 2014	7	7	7	NA

Note: NA = Not Available, PCP=Precipitation, Tmax=Maximum Temperature, Tmin=Minimum Temperature,, Kibremengist official name is Adola

The Mean Absolute Error (MAE), Root Mean Square Error (RMSE), PBIAS, and Coefficient of Determination (R2) were used to compare the CHRIPS and CHIRTS data to the ground truth. For example, Table S2 shows that the sample station for Delomena has an R2 for rainfall of 0.604 and 0.757 for mean temperature.

Table S2. Comparison metrics for station data with CHRIPS and CHIRTS

Station	Type	Bias	MAE	RMSE	R2
Delomena	Precipitation	-0.156562	1.856976	2.882518	0.604034
Delomena	Mean Temperature	0.134965	1.045273	1.342489	0.756614
Kibremengist	Precipitation	0.542582	1.883854	2.630695	0.437942

Kibremengist	Mean Temperature	0.142520	1.168628	1.520184	0.824438
Bore	Temperature	-1.859474	2.023267	2.374163	0.634567

Temperature was more strongly correlated than precipitation when we examined the station's correlation with satellite CHRIPS and CHRIRTS data, as indicated in Table 2. The Grubbs test for outlier detection (Adikaram et al., 2015), NSTH, Pettit (Pettitt, 1979), and Buishand tests (Buishand, 1982) were used to test rainfall data homogeneity and identify changes in the time series data. The multiple imputation method was used to fill in the missing temperature data. The Normal-ratio method was used to estimate missing data when rainfall differs by more than 10% and when stations are unevenly distributed.

$$\mathrm{PX}=\frac{1}{N}\left(\frac{NX}{NA}PA+\frac{NX}{NB}PB+\frac{NX}{NC}PC+\_\_\_\_\_\_+\frac{NX}{NN}\right) \quad 1$$

Where PX is the rainfall for the place with the missing record, N is the number of stations used in the computation, PA, PB, PC,… PN is the corresponding precipitation at the index stations, and Nx is the normal annual precipitation of station X. NA, NB, and NC are the normal annual precipitation of the surrounding stations. Hydrological data from the Chenemsa gauge station gathered from MoWE between 1988 and 2011 were used to validate and calibrate the QSWAT+ simulation. Table S3 provides a summary of the sources and data types.

Table S3. Summarized data, type, software, and sources used in the study

No	Data Type and Tools	Sources
1	Meteorological data	Ethiopian National Meteorological Institute (ENMI)
2	Hydrological data	Ministry of Water and Energy (MoWE)
3	Soil data(map)	HWSD version 1.2
4	DEM	https://srtm.csi.cgiar.org/
5	LULC data	QGIS MapSWAT V3.0 plugin, Copernicus Global land cover 2009
6	ArcGIS 10.4.1 Software	Download
7	CROPWAT8.0 Software's	FAO
8	Climate Change Data CMIP 6	Spatial resolution 0.27°(30×30km) downloaded from NASANCCS-NEX-GDDPCMIP6 using GEE, Online free access

1.2 Climate Model Selection and Downscaling

This study analyzed historical data from 1985 to 2014 using the Global Climate Model (GCM) and future scenarios from 2015 to 2074. Using observational data from five stations and the CMIP6 GCM, this study assessed the impact of climate change on surface water availability. To select a climate model, the Comprehensive Ranking Index (CRI) was used to evaluate the performance of 22 global climate models from the NASA/GDDP-CMIP6 Earth Engine Image collection against the CHIRPS and CHIRTS observational datasets (Verdin et al., 2020). The study assessed GCMs' ability to mimic past climate variability using correlation coefficients, bias, and RMSE metrics, identifying proficient models using a single ranking system (Figure 2). The Python scripts for evaluating and selecting the CMIP6 models are easily accessible through Zenodo (Ali, 2026). Accordingly, the study utilizes the Canadian Earth System Model version 5(CanESM5), ranked second as per the CRI evaluation (Figure 2) for temperature and the ensemble of the six models for precipitation. It is an improved version of the Canadian Centre for Climate Modelling and Analysis (Swart et al., 2019). Significant research has been conducted in Ethiopia using CanESM2, which is the previous version of CanESM5 (Adgolign et al., 2015; Asmerom, 2015; Kebede, 2013; Negewo, 2020b)

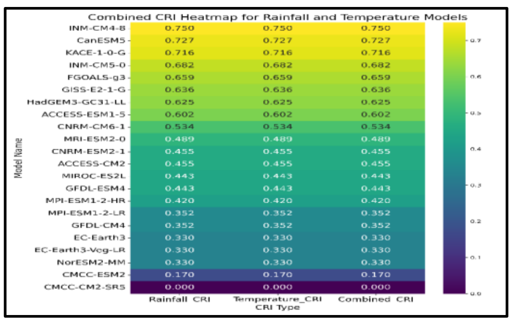
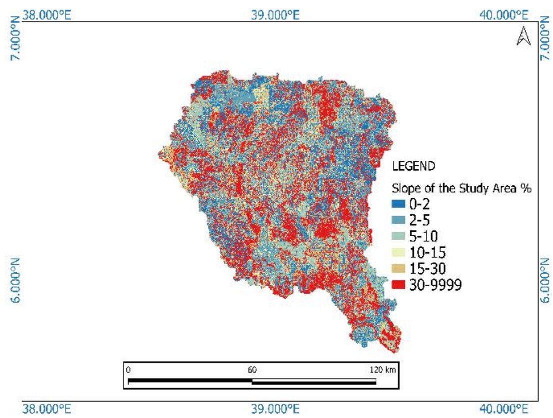


Figure 1. Combined CRI for Rainfall and Temperature Models (Higher Better)

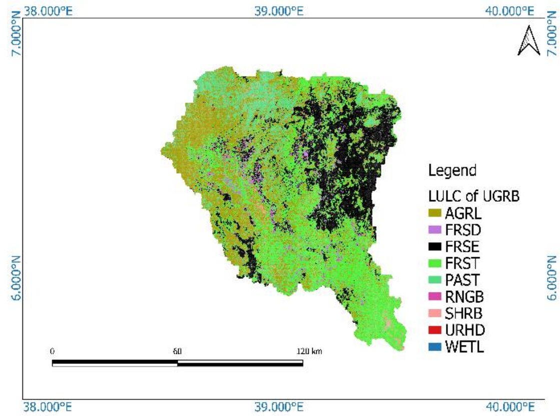
The study utilized temperature and rainfall change projections from CMIP6 Global Climate Models, which were downloaded, and bias correction and downscaling were performed for five Upper Ganale Basin stations using SD-GCM V2 software. SD-GCM V2 is a decision-support tool for bias correction and downscaling of CMIP6 data, using three methods to assess local climate change impacts using a robust statistical downscaling technique (SD-GCM, 2017). The Delta, Empirical Quantile Mapping (EQM), and Quantile Mapping (QM) methods were compared for temperature and precipitation in SD-GCM V2, and the best method was selected using statistical analysis. The entire process is described in a previously published study (Ali E. A et al, 2026).

1.3 Spatial Data Analysis

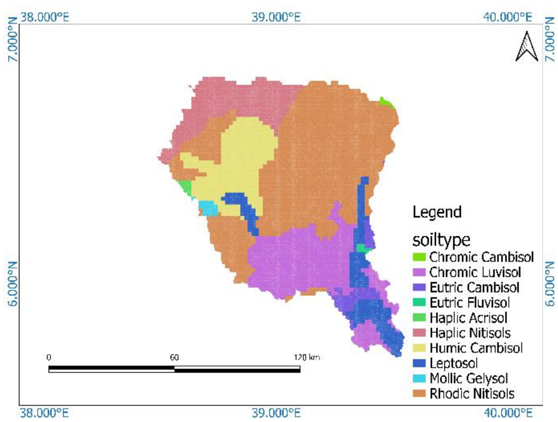
The DEM was used to calculate the land slope in the UGRB, showing that 82.6% of the region has a slope of over 10%, with 17.4% of the sub-basins having slopes between 0 and 10%.Figure 3.a).



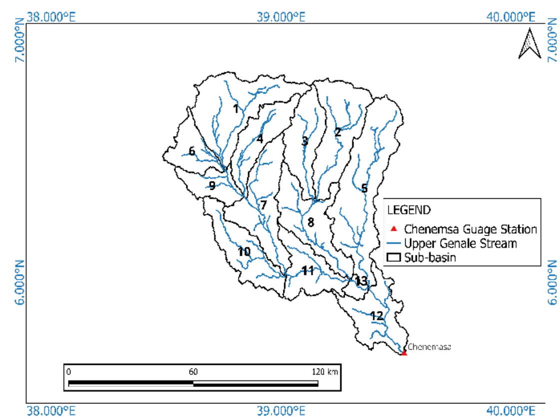
(a)



(b)



(c)



(d)

Figure 2.(a) Slope of the study area in percent, (b) LULC of the study area, (c) Soil type of the study area, (d) Sub-basin of the Study Area, Chenemsa hydrological gauge station, Upper Ganale Stream

The classification of the LULC maps was performed using the Google Earth Engine (GEE) tool, a rapid analysis platform that leverages Google's processing power (Z. Zhao et al., 2024). Landsat satellite images from Landsat-5, Landsat-7, and Landsat-8 were used for the classification process(Xie et al., 2019). Accordingly, the land use and land cover of the area in 2009 were mostly mixed forests, agriculture, and evergreen forests, in that order, with dynamic changes in land use and land cover (Table S4 and Figure 3b).

Table S4. Land Use and Land Cover of Upper Genale Sub-Basin of 2009 GCC

No	Land use/cover	Area Coverage (ha)	Area Coverage %
1	Agricultural Land	2502661.68	27.42
2	Pasture	92,758.55	10.16
3	Urban and homestead,	874.2	0.1
4	Evergreen Forest	192,124.33	21.05
5	Forest Mixed	319,668.61	35.02

6	Range, Grass, and Shrubs	57,071.96,	6.25
	Total	912,791.82	100

This study used a soil map and the QSWAT+ model to analyze hydrological parameters. The soil shapefile was obtained from the MoWE and extracted using QGIS 3.40. The soil classification results revealed that Nitosols, Luvisols, Cambisols, and Leptosols accounted for 99% of the sub-basin as indicated in Table S5 and Figure 3c.

Table S5. Soil Type of the Upper Genale Sub-Basin

Soil Type	Area in Ha	%
Acrisol	3023.71	0.331
Cambisol	135151.56	14.806
Fluvisol	2021.81	0.221
Gelysol	5816.51	0.637
Leptosol	70315.01	7.703
Luvisol	164308.68	18.000
Nitisols	532154.53	58.300
Total	912791.81	100

1.4 QSWAT+ Tool Modelling Approach

QSWAT+ is an open-source QGIS interface for SWAT+, which is an object-based successor to SWAT. The SWAT+ is a physically based soil and water assessment tool developed by the United States Department of Agriculture Research Services to simulate the impact of climate change on river flow (Arnold et al., 1998). The SWAT model simulates evapotranspiration, surface runoff, infiltration, percolation, shallow and deep aquifer flows, and channel routing. Runoff volume was estimated from rainfall data using the modified Soil Conservation Service (SCS) curve number method (Equation (2)) for each Hydrologic Response Unit:

$$Qsurf = \frac{(Rday - 0.2S)^2}{(Rday + 0.8S)} \qquad 2$$

where Qsurf is the accumulated daily surface runoff (mm), Rday is the daily rainfall (mm), and S is the retention parameter. In this method, the curve number varies nonlinearly with the moisture content of the soil profile, reaching its lowest value when the soil profile moisture content approaches the wilting point and increasing to nearly 100 as the soil approaches saturation. Parameter S is related to the curve number and was estimated using Equation (3):

$$S = \frac{25400 - 254CN}{CN} \qquad 3$$

where CN is the curve number for each day.

The hydrologic routines simulated by SWAT are based on the following water balance equation (Equation (4):

$$SWt = SWo + \sum_{i=1}^{t}(Rday - Qsur - Ea - Wseep - Qgw) \qquad 4$$

where SWti s the final soil water content (mm), SWo is the initial soil water content (mm), t is time (days), Rday is the amount of precipitation on day i (mm), Qsurf is the amount of surface runoff on day i (mm), Ea is the amount of evapotranspiration on day i (mm), Wseep is the amount of water entering the vadose zone from the soil profile on day i (mm), and Qgw is the amount of return flow on day i (mm). For this study, the FAO Penman-Monteith method was used for ET estimation, and variable methods were used for flood routing.

1.4.1 Input File Preparation, Model Setup, and Simulation

The process involved input data preparation using weather data tools, including climate, soil, hydrological, and topographic data. Projections were made using QGIS 3.40.11, and the DEM was then loaded. The model was adjusted to the location of the gauging station for calibration. Stream networks, sub-watersheds, and topographic parameters were calculated using tools, resulting in 13 sub-basins (Figure 3d). The Upper Genale River sub-basin was divided into 731 HRUs based on the threshold values for the LULC, soil, and slope categories, ensuring 100% coverage of their respective areas, as per the SWAT+ user's manual (Figure 4).

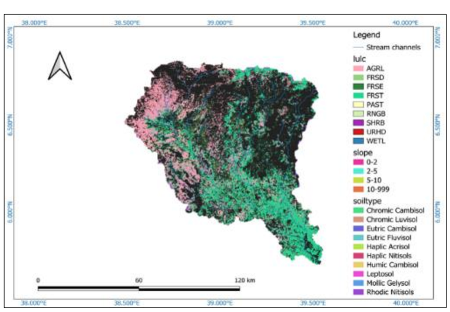


Figure 3. Upper Ganale River sub-basin Hydrological Response Units with slope, soil, and LULC layers

1.4.2 Model Calibration and Validation

The Chinemsa hydrological station was used for both calibration and validation. With a three-year (1985–1987) warm-up period, data from 1988–2002 were used for calibration, and data from 2003–2011 were used for validation. A sensitivity analysis was performed using the Sobol method. This is a global sensitivity method that explores the entire parameter space using Monte Carlo sampling (Bauwens, 2012; Pianosi et al., 2016; Umali & Lansigan, 2020).

1.4.3 Model Performance Evaluation

The model performance was evaluated to compare the model simulation outputs with the observed data. Consequently, the model performance was assessed using the coefficient of determination (R2), Nash and Sutcliffe simulation efficiency (NSE), and percent bias (PBIAS). The Nash-Sutcliffe efficiency (NSE) is a statistical measure that quantifies the proportion of residual variance in a sample compared to the measured data variance (Nash and Sutcliffe, 1972). The NSE can be computed using the following equation:

$$\text{NSE}= 1-\frac{\sum_{i=1}^{m}(Qoi-Qsi)2}{\sum_{i=1}^{m}(Qoi-Qoav)2} \qquad 5$$

The coefficient of determination (R2) is the measurement of the correlation between simulated and observed values, known as the coefficient of determination. The R2 varies from 0 to 1, where an approaching 1 indicates better performance. R2 can be calculated as follows:

$$\text{R2}\ \frac{\sum_{i=1}^{m}\left((Qoi-Qoav)(Qsi-Qsav)\right)2}{\sum_{i=1}^{m}(Qoi-Qoav)2\sum_{i=1}^{m}(Qsi-Qsav)2} \qquad 6$$

The percent bias (PBIAS) calculates the average tendency of simulated values to be larger or smaller than observed values. A PBIAS with a smaller magnitude indicates a better model performance. PBIAS can be calculated as follows:

$$\text{PBIAS}=100 * \sum_{i=m}^{m}\left(\left(\frac{Qoi-QSi}{QSi}\right)^{2}\right) \qquad 7$$

where m is the total number of observations, Qoi and Qsi are the observed and simulated discharge at the ith observation, respectively, and Q0av and Qsav are the mean observed and simulated data over the observed and simulation periods, respectively. Where Q is the discharge, and the optimal value of PBIAS is zero. A positive PBIAS value indicates an underestimation by the model, and negative values represent overestimation.

1.4.4 Sensitivity Analysis

A sensitivity analysis was performed to identify and assess the most sensitive parameters of the watershed under investigation. The parameters included in the sensitivity analysis were those that affect the values of watershed PET, runoff, groundwater flow, a routing process, reservoir/lake/pond, water abstractions for domestic water supply, irrigation schemes, subbasin, soil, and HRU. Fifteen parameters from the literature (Shigute et al., 2022) were selected to run the sensitivity analysis of the parameters that affect the watersheds. Using the SWAT+ Toolbox program, the Sobol method, global sensitivity analysis, and mapping out the sensitive parameters (Jiménez-Navarro et al., 2024).
The Parameter importance was evaluated based on the relative sensitivity magnitude and ranking. Statistical significance testing (t-statistic and p-value) was not performed because the SWAT+ Toolbox does not employ a regression-based sensitivity analysis. The Sobol global sensitivity analysis embedded in the SWAT+ Toolbox was used to determine the most significant flow parameters. The sensitivity of streamflow was determined using average daily streamflow data for 24 years (1988–2011) recorded at the Chenemsa Hydrological Station.

1.4.5 Analysis of Variability of Water Availability

A comprehensive hydro-climatic data analysis was conducted to assess the variability and trends in water availability in the UGRB. The goal was to analyze, quantify, and validate temporal variations in rainfall and streamflow. The statistical tools used were the Coefficient of Variation (CV), Mann-Kendall (MK) test, and Sen's slope estimator.

Coefficient of Variation (CV): CV is a statistical measure that shows how data vary around the mean; it is the ratio of the standard deviation to the mean. This study classified variability as follows:

CV < 20% → Low variability (normal); 20% ≤ CV ≤ 30% → Moderate variability; CV > 30% → High variability. In previous studies, the CV was used to characterize rainfall variability (ALI et al., 2025; Bedada et al., 2025).

$$CV = \frac{\sigma}{X} \tag{8}$$

where σ and X denote the standard deviation and mean of rainfall, respectively

Mann-Kendall (MK) Test: The Mann-Kendall (MK) test is a popular nonparametric trend test that is unaffected by outliers and missing values and is independent of data distribution (Hamed & Rao, 2018). Before the Mann-Kendall trend test, lag1-autocorrelation was examined for rainfall and streamflow over annual and seasonal time periods because serial autocorrelation increases the probability of the MK trend test. In this study, the significance of the serial correlation was detected using Student's t-test at lag-1 for the streamflow data series, with the following formula:

$$t=\rho 1\sqrt{\frac{n-2}{1-\rho^2}} \quad 9$$

where the test statistic t has a Student's t distribution with (n-2) degrees of freedom. If $|t| \geq t\ \alpha/2$, the null hypothesis about serial independence is rejected at significance level α, i.e., significant autocorrelation

For stations with significant autocorrelation, the trend was tested and confirmed using a linear regression. After analyzing the hydro-climatic trends of the stations, the average duration of the records at the same stations was chosen to investigate the relationship between their trends using correlation. The test statistic S is calculated as follows.

$$S=\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} sgn(Xj-Xi) \quad 10$$

where Xj denotes the consecutive data values, n denotes the number of the data set, and such that j > i, and where the sgn function is given as

$$sgn(\theta)=\begin{cases} 1 & if\ \theta>0 \\ 0 & if\ \theta=0 \\ -1 & if\ \theta<0 \end{cases} \quad 11$$

The variance, Var(s), was calculated by

$$Var(s)=\frac{n(n+1)(2n+5)-\sum_{i=1}^{m} ti(ti-1)(2t+5)}{18} \quad 12$$

where n denotes the number of observations, m is the number of tied groups, and ti is the number of ties of the sample at time i.

The standard test statistic Zs was calculated using:

$$Zs=\begin{cases} \frac{S-1}{\sqrt{Var(S)}} & S>0 \\ 0 & S=0 \\ \frac{S+1}{\sqrt{Var(S)}} & S<0 \end{cases} \quad 13$$

In a Z test, the null hypothesis implies no trend, whereas the alternative hypothesis means a significant change has occurred. The values of Zs > 0 and Zs < 0 indicate an increasing and decreasing trend, respectively. Moreover, a significant trend at 0.05 and 0.01 significance levels exists when the /Z/ > 1.96

and /Z/ > 2.58, respectively. In this study, 0.05 and 0.01 significance levels were considered to analyze the hypothesis.

The significance levels (p-values) for each trend test can be obtained from the relationship given as

$$p=0.5-\varphi(|Zmk|) \qquad 14$$

where φ () denotes the cumulative distribution function (CDF) of a standard normal variate. At a significance level of 5%, if $p \leq 0.05$, then the existing trend is considered to be statistically significant.

Sen's slope estimator is a non-parametric analysis to estimate the magnitude and direction of the trend in a sample of N pairs of data. The slope (Qi) was calculated by

$$Qi=\frac{Xj-Xk}{j-k}\,(i=1,\ldots,N) \qquad 15$$

where xj is the data value at j and xk represents the data value at k (j > k). When there is one data point in each period, then N=n(n-1)/2, where n is the number of periods. But if there are more data points, then N < n(n-1)/2, where n is the total number of observations. The values of N are arranged from smallest to largest. Then, the median of the slope or Sen's slope estimator (Qmid) was calculated by

$$Qmid=\begin{cases}\frac{Q(N+1)}{2}, & if\ N\ is\ odd\\ \frac{Q\left(\frac{N}{2}\right)+Q(N+2)/2}{2} & if\ N\ is\ even\end{cases} \qquad 16$$

When the value *Qmid* > 0, it indicates an increasing trend, while *Qmid* < 0 represents a decreasing trend.

1.4.6 Assessing the climate change impact on Stream flow and Water Balance

A baseline period (1985-2014) and two time windows (near-term (2015-2044), which is a near climate change extended scenario (SSP2-4.5, SSP5-8.5), and mid-term (2045-2074), which is a mid-term climate change extended scenario (SSP2-4.5, SSP5-8.5) were taken into account. The study used the following equation to calculate temperature, precipitation, and stream flow changes.

Percent change (%) of water balance and stream flow

The percent change in future water balance and stream flow impacts analyzed for each SSP climate scenario of each climate model for each that averaged over the whole sub-basin of each in the Basin predicted two-time windows (near term and mid term) from their corresponding baseline period (1985-2014) was calculated by the following equation.

%Change=100× | (WBf-WB), monthly, seasonally, annually/WB, monthly, seasonal, annually 17

Where %change is the percent change of each water balance component and stream flow, WB is the average water balance components and stream flow of the current (baseline) period (1985-2014), WBf average water balance component and stream flow of future of near term period (2015-2044), mid-term of the period (2045-2074) of each SSP Scenarios (SSP2-4.5 and SSP5-8.5). WB represents the Water balance components (rainfall, surface runoff, lateral flow, percolation of water, and water yield).

1.4.7 Machine Learning, Hybrid Model Design, and Analysis Approach

The study employs a hybrid data-driven and process-based modeling framework for streamflow forecasting (Belina et al., 2025a; Kim et al., 2019; W. Zhao et al., 2025). It combines the physically based QSWAT+ model with advanced machine learning algorithms such as Random Forest (RF), XGBoost, LightGBM, and Long Short-Term Memory (LSTM).

1.4.8 Data Integration and Standardization

For this study, the five meteorological station data used in the QSWAT+ simulation were also input to the four selected models. Thus, station mapping was carried out using numerical identifiers. Date temporal alignment was converted to a uniform date-time format to enable precise merging of climate variables (pcp, tmp, hmd, wnd, slr) with QSWAT+ simulations and observed flow records. All historical climate data were loaded from 1985 to 2014. Chenemsa station daily runoff for the 1988 to 2011 period was used with meteorological forcing data. Similarly, gridded data of reanalysis of the selected CMIP6 models were also incorporated for spatial representation.

To convert unprocessed climatic and hydrological data into a format suitable for supervised machine learning and deep learning (LSTM), feature engineering is essential. It makes it possible to extract watershed memory and seasonal sensitivity from the available data. Lagged features, time-based, and rolling window statistics were used as an engineering feature process. Accordingly, 1-day and 7-day lags for precipitation, QSWAT-PLUS simulated flow, and observed stream flow were used to account for temporal dependencies and determine the most current weekly trends or longer-duration recession features. Similarly, a 7-day rolling mean was used to calculate 7-day rolling means for flow and precipitation in order to capture short-term environmental trends. The year, month, and day-of-year components were extracted to obtain temporal features and produce periodic seasonal signals. Seasonality is a primary driver in hydrological modeling.

Finally, data standardization through utilizing standard scaling was carried out to remove the mean and scale features to unit variance to ensure that features with large ranges (like flow) do not dominate those with smaller ranges (like humidity) during model training. Specifically, the LSTM architecture was reshaped from a 2D array (samples, features) into a 3D tensor (samples, time steps, features) and ensured each sample is treated as a single time step (time steps=1) containing all engineered features. In general, delays and rolling windows allow for the transformation of a collection of static climatic records into a dynamic time-series structure. This allows the models to grasp not only the immediate relationship between weather and flow but also the watershed's memory, which is critical for accurate streamflow forecasting.

1.4.9 Machine Learning Model Description and Training

The forecasting framework employs four distinct machine-learning architectures, each leveraging specialized libraries and optimization methods. The Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) is an extension of the recurrent neural network (RNN) that was created to solve the problem of vanishing and exploding gradients in RNNs. The three main gates of LSTM are the input gate, the forget gate, and the output gate. To capture temporal dependencies, this deep learning model was built with Tensor Flow/Keras. In this study, an Adam optimizer was used to train a hidden layer with 50 units and a ReLU activation function over 50 epochs.

Random Forest (RF) is a tree-based supervised machine-learning model developed with scikit-learn. This ensemble of decision trees decreases variance via bagging. It was trained with 100 estimators to detect non-linear hydrological correlations. The other one is Extreme Gradient Boosting (XGBoost), which uses the xgboost package. This gradient boosting framework optimizes a squared error objective function (reg: squared error). It creates trees consecutively to reduce residuals from earlier iterations. The Light Gradient Boosting Machine (LightGBM), implemented via the lightgbm package, leverages leaf-wise tree development for improved efficiency and faster training, while also targeting Mean Squared Error (MSE).

1.4.10 Hybrid Stacking Ensemble Architecture

The modeling framework employs a two-tier stacking architecture, with base models as level 0 and a meta learner as level 1. At the base level, diverse predictive signals are generated from four machine learning models (Random Forest, XGBoost, LightGBM, and LSTM) and the physically based QSWAT+ model. At level 1 (Meta-Learner), a Ridge Regression model integrates these predictions. By using L2 regularization, the Ridge meta-learner optimally weights the base models to reduce systematic bias while preventing overfitting. Ridge Regression was used as the meta-learner to combine outputs from QSWATPLUS and machine-learning models, minimizing multicollinearity and overfitting. Lin et al. (2022) in Remote Sensing and Chen et al. 2024) in Water Supply have validated similar methodologies, demonstrating improved runoff and drought forecast accuracy through ridge-based ensemble integration. In this study, Bayesian

Median Analysis of the weighting method was used. Weights are assigned dynamically based on the R2 performance of individual models during the validation period.

1.4.11 Statistical Evaluation and Residual Analysis

Model performance metrics were evaluated for the 1988-2011 period using NSE, R2, and PBIAS (formulas shown in equations 5-7). Similarly, the Mann-Kendall and Sen's Slope tests were used to assess significant monotonic trends on daily and seasonal scales. Finally, the Rescaled Adjusted Partial Sums (RAPS) analysis (Alexandersson, 1986) was used to identify mutation points in the flow series. The Rescaled Adjusted Partial Sums (RAPS) method is a non-parametric statistical technique used to identify intrinsic fluctuation patterns, mutation points, and persistent changes within a time series. It quantifies the cumulative deviation of the time series from its mean, scaled by its standard deviation. This allows for visual detection of trends, shifts, and cycles.

The calculation formula for the RAPS sequence is given by:

$$Rk = \sum_{t=1}^{k} \frac{Xt - X}{Sn} \qquad 18$$

Where: Xt is the value of the time series at time(t), X is the mean of the entire time series, Sn is the standard deviation of the entire time series, k is the index of the RAPS sequence, ranging from 1 to the total number of data points n.

1.4.12 Future Projections (CMIP6 Integration) Scenario Analysis

Integrated ensemble of climate model data under SSP245 (Moderate) and SSP585 (High) scenarios for near-term (2015-2044) and mid-century (2045-2074) windows was loaded to be used by the model. Finally, iterative forecasting using a bootstrapping loop, in which previous predictions serve as inputs for subsequent steps, was employed, enabling autonomous 30-year streamflow simulations for each scenario.

The selection of machine learning algorithms, including a hybrid, followed the framework described by Belina et al. (2025a), Kim et al. (2019), and Li et al. (2024), who demonstrated the effectiveness of RF, LightGBM, XGBoost, CatBoost, LSTM, and MLP for streamflow prediction in Ethiopia and other regions. In line with their findings, the hybrid model was designed to improve the performance of QSWATPLUS streamflow prediction. To operationalize the model, the dataset covering 1988–2011 was compiled as a daily series of input and output variables (n = 5472). These were then stochastically partitioned into training (70%) and testing (30%) (Bojer et al., 2024). The model was used to forecast streamflow from historical data and future projections (2015–2074) based on the ensemble mean of the six selected climate model data under the SSP2-4.5 and SSP5-8.5 scenarios. The Python code for data preparation, model training, and evaluation is openly available on Zenodo (ALI, 2026).

2 Result

Table S5. Sensitivity analysis results of the parameters

Ranking	Sensitivity	Name	Best Sensitivity Value	Change Type	ID
1	0.91397	cn3_swf	0.9057	Replace	7
2	0.52965	perco	0.928	Replace	2
3	0.06593	cn2	-12.393	Replace	10
4	0.06510	canmx	1	Percent	1
5	0.06078	revap_min	20.9454	Replace	4
6	0.05863	revap_co	4.6875	Replace	11
7	0.00512	alpha	9.3750	Percent	9
8	0.00054	epco	19.0625	Percent	6
9	0.00044	chk	0.8719	Replace	12
10	0.00000	bf_max	-6.2500	Percent	15
11	-0.00000	lat_ttime	0.7609	Replace	8
12	-0.00019	flo_min	0.4688	Percent	3
13	-0.00135	awc	3.1250	Percent	13
14	-0.00376	esco	9.6875	Percent	14
15	-0.00915	k	0.2812	Replace	5

Table S6. Best fitted parameter values under Automatic Calibration by Latin hypercube Sampling Iterations (CALSI) algorithm

ID	Name	Group	Change Type	Fit Value	Set Min	Set Max
1	cn2(SCS curve numbers)	hru	Percent	-12.393	-20	20
2	Esco (Soil Evaporation Compensation Factor)	hru	Replace	0.326	0	1
3	Epco (plant uptake compensation factor)	hru	Percent	0.6717	0	1
4	Canmx (Maximum canopy storage)	hru	Replace	1	1	100
5	Perco(percolation coefficient)	hru	Replace	0.9280	0	1

6	Lat-ttime(Lateral flow travel time(days))	hru	Percent	3.7958	-10	20
7	cn3_swf(pothole evaporation coefficient)	hru	Replace	0.9057	0	1
8	Alpha(base flow alpha factor)	aqu	Replace	0.2099	0.01	0.9
9	Flo-min(minimum aquifer storage to allow return flow)	aqu	Percent	15.2005	0	20
10	Revap-co(groundwater revap coefficient)	aqu	Replace	0.1055	0.02	0.2
11	Revap-min(the threshold depth of water in the shallow aquifer for revap or percolation to the deep aquifer to occur	aqu	Replace	20.9454	0	50
12	Bf-max(the maximum baseflow, the deep aquifer's gift to rivers)	aqu	Replace	1.4082	0.1	2
13	Awc(Available water capacity of the soil layers)	sol	Percent	1.7833	-10	10
14	K(Saturated Hydraulic conductivity)	sol	Percent	11.4619	-10	20
15	Chk (the channel's keeper, shaping flow within riverbeds)	rte	Percent	4.4030	-20	20

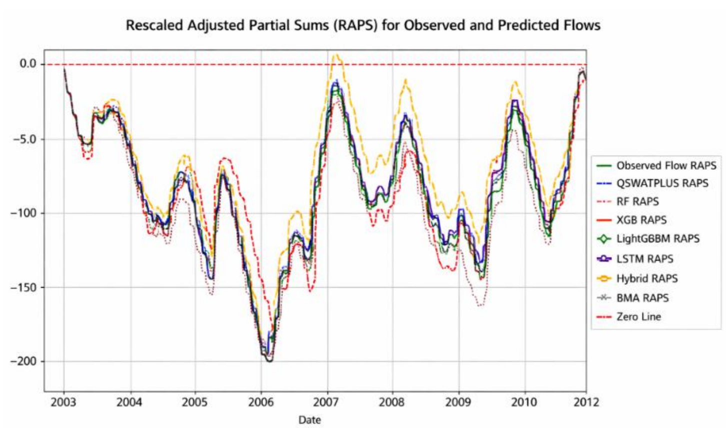


Figure 4. Rescaled Adjusted Partial Sum for Observed and Predicted Streamflow from 2003-2001

1.1 Future Rainfall and Temperature Projection

1.1.1 Future Temperature Projection

Meteorological stations like Delomena(DLM), Hagereselam (HGRS), Bore (BOR), Adola (Kibre Mengist)(KBRM), and Negele(NGL) were used to investigate the trends and variability of mean annual and seasonal minimum and maximum temperatures (Table 11). The DLM station recorded the highest mean annual maximum temperature, which varied between 18.67 and 28.93 °C. The average annual minimum temperature ranged from 7.13 to 12.98 °C, with the lowest value recorded at the HGRS station. The annual

CV and SD of the mean maximum temperature ranged from 1.362 to 3.76 percent and 0.337 to 0.756 °C, respectively. Similarly, the annual CV and SD of the Tmin range from 3.64 to 22.657 percent and 0.471 to 1.621 °C, respectively (Table 11).

Seasonally, the maximum mean temperature weighted average for spring (MAM), summer (JJA), and autumn (SON) ranged from 22.86 to 25.55 °C, with CV ranging from 2.85 to 4.23 percent, respectively. The Winter (DJF) period's highest recorded maximum temperature, which is in line with the finding of Dandesa et al. 2017), who found that the Borena Zone had the highest maximum temperature during DJF. The weighted average mean seasonal minimum temperature varied between 11.8 and 12.6 °C, with CV ranging from 9.36 to 12.6 percent, respectively, as shown in Table 11. HGRS and DLM stations had the lowest and highest temperatures over the course of the seasons, respectively.

Table 6. Annual and Seasonal Tmax and Tmin trend analysis and mean and CV in percentages for selected meteorological stations in the UGRB for the base period (1985-2014)

Season	Parameters	Maximum Temperature(Tmax)					Mean	Minimum Temperature(Tmin)					Mean
		DLM	KBRM	NGL	BOR	HGRS		DLM	KBRM	NGL	BOR	HGRS	
Annual	Slope	-0.01	0.02	0.07	0.04	-0.03		0.04	0.06	0.07	0.02	0.05	
	P	0.86	0.03*	0.0*	0.0*	0.03*		0.0*	0*	0*	0.1	0*	
	Z	-0.18	2.21	5.46	5.32	-2.2		3.03	3.93	5.32		4.25	
	Mean	28.9	25.90	26.5	24.8	18.7	24.95	12.98	12.8	15.5	10.5	7.15	11.8
	CV	2.31	2.92	2.59	1.36	3.76	2.59	3.64	3.69	7.74	4.64	22.66	8.47
	SD	0.67	0.76	0.69	0.34	0.70	0.63	0.48	0.47	1.20	0.49	1.62	0.85
Spring	Z	-0.71	0.64	4.28	3.18	-1.39		2.93	3.6	3.96	0.21	2.78	
	P	0.48	0.52	0.0*	0.0*	0.16		0.0*	0.0*	0*	0.15	0.01*	
	Slope	-0.01	0.01	0.09	0.03	-0.03		0.05	0.07	0.07	0.02	0.04	
	Mean	29.45	26.91	26.7	25.5	19.3	25.55	13.8	13.7	16.3	11.5	7.88	12.6

	CV	3.80	3.51	3.71	2.41	3.52	3.39	4.38	6.92	8.51	4.45	22.94	9.44
Summer	Z	0.1	2.85	4.78	4.14	-1.7		1.96	2.75	3.96	0.15	2.93	
	P	0.91	0.00*	0.0*	0.0*	0.09		0.0*	0.01*	0*	0.25	0.00*	
	Slope	0.002	0.03	0.07	0.05	-0.04		3.03	3.93	5.32		4.25	
	Mean	27.17	23.00	24.1	23.2	16.9	22.86	13.0	13.8	14.8	11.9	7.14	12.1
	CV	3.43	4.34	4.66	3.49	5.24	4.23	10.0	4.17	6.61	4.95	21.22	9.39
Autumn	Z	1.8	3	4.8	5.1	-1.5		*2.93*	*3.6*	*3.96*	*0.21*	*2.78*	
	P	0.08	0.00*	0.0*	0.0*	0.13		0.04	0.06	0.07	0.02	0.05	
	Slope	0.02	0.03	0.07	0.05	-0.03		0.0*	0*	0*	0.06	0*	
	Mean	28.28	25.51	26.1	23.99	18.1	24.41	13.1	13.1	15.5	10.2	7.06	11.8
	CV	2.02	3.08	2.09	2.07	4.99	2.85	4.65	4.32	6.77	5.38	25.69	9.36

The base period result of the MK analysis on mean annual Tmax showed a statistically significant increasing trend at 0.05 for KBRM, and for NGL and BOR stations at 0.01 levels of significance (Table 11). The HGRS stations showed decreasing trends at a 0.05 significant level, whereas the DLM station showed no trend in annual mean Tmax. Similarly, all stations except BOR showed a statistically increasing trend at a 0.01 level of significance. From the result of the annual mean trend analysis, there is a greater increasing trend for most stations with respect to Tmin than Tmax. The seasonal trend for spring, summer, and autumn over the stations showed an increasing trend for all stations for Tmin, except BOR, which showed an insignificant level at all. Regarding Tmax, NGL, and BOR showed a significant level of increasing trend across the three seasons, whereas KBRM only showed a significant increase at a 0.01 level in summer and autumn (Table 11).

The result of future maximum temperature (Tmax) and minimum temperature (Tmin) for mean annual and seasonal change in temperature under SSP 2-4.5 and SSP5-8.5 during the near and midterm period was indicated in Table 12. Accordingly, Tmax increased by 0.51 °C, Tmin increased by 0.20 °C under SSP2 4.5 near term, whereas Tmax increased by 0.22 °C, Tmin increased by 0.19 °C under SSP5 8.5 near term. Under midterm scenarios, Tmax increased by 0.796 °C, Tmin increased by 1.049 °C under SSP2 4.5, whereas Tmax increased by 1.049 °C, Tmin increased by 2.055 °C under SSP5 8.5. Compared to the Near Term, warming is noticeably more significant in the midterm in this scenario under SSP2-4.5 (Table 12).

Compared with SSP2-4.5 for the Mid Term, the predicted warming is significantly higher. Thus, the highest maximum and minimum annual temperatures under SSP5-8.5 in the Mid Term, respectively, indicate increases of 0.3 °C and 0.331 °C per decade, suggesting that the future will be warmer than the base year.

Table 7. Mean annual and Seasonal temperature change in °C under different scenarios

Scenario	*SSP2-4.5*				*SSP5-8.5*			
Timeframe	*Near term*		*Midterm*		*Near term*		*Mid term*	
Variables	*Tmax*	*Tminx*	*Tmax*	*Tmin*	*Tmax*	*Tmin*	*Tmax*	*Tmin*
Annual	*0.514*	*0.201*	*0.796*	*1.049*	*0.218*	*0.193*	*1.049*	*2.055*
Autumn	0.190	-0.392	0.633	0.439	0.063	-0.423	1.406	1.624
Winter	0.002	-0.128	-0.484	0.545	-0.283	-0.060	-0.239	1.88
Spring	0.603	-0.240	1.462	0.808	0.370	-0.192	1.319	1.461
Summer	1.363	0.553	1.782	1.311	0.732	0.364	2.227	2.209

Regarding the result of seasonal temperature change, there are noticeable seasonal differences: summer shows the strongest warming, where Tmax increased by 1.78 °C under SSP2-4.5 and by 2.23 °C under SSP5-8.5 mid-term, respectively. Spring shows significant Tmax increases by 1.462°C under SSP2-4.5 and increases by 1.32°C under SSP5-8.5 scenarios. Autumn shows moderate warming, and winter shows mixed signals, including near-term Tmin cooling but sharp mid-term increases under SSP5-8.5 by 1.88 °C (Table 12). The result of the study is consistent with the findings of East African CMIP6 assessments, which project 0.5-1.0 °C warming by 2040 and 1.5-2.5 °C by 2060 under SSP2-4.5, with stronger signals under SSP5-8.5 (Gebrechorkos et al., 2019; IPCC, 2021).

Table 8. Summary Result of projected mean, Tmin, and Trend under SSP2-4.5 and SSP5-8.5 Scenarios for Annual and Seasonal basis

Scenario	SSP2-4.5						SSP5-8.5					
Timeframe	Near term			Midterm			Near term			Mid term		
Variables	Tmax			Tmin			Tmax			Tmin		
Trend	P	Z	S	P	Z	S	P	Z	S	P	Z	S
Annual	0.03*	2.20	0.01	0.80	0.00	6.97	0.00*	5.48	0.03	0.00**	9.55	0.06

Autumn	0.003**	2.95	0.01	0.00**	6.83	0.03	0.00**	7.56	0.04	0.00**	9.52	0.07
Spring	0.002**	3.07	0.02	0.00**	6.27	0.03	0.37	0.00	6.18	0.00**	8.82	0.06
Summer	0.14	1.46	0.01	0.00**	5.95	0.02	0.00*	6.18	0.05	0.00**	8.82	0.06

NB ** and * show significance at $\alpha = 0.01$ and 0.05 respectively

The future projected results summary of the MK analysis on mean annual Tmax and Tmin in the Upper Genale Basin is shown in Table 13. The results showed a statistically significant increasing trend for both Tmax and Tmin across both scenarios at 0.05 on an annual basis for all periods except the Midterm SSP2-4.5 period, which showed no trend for Tmin. Regarding seasonal MK trends, both Tmax and Tmin showed an increasing trend at a significant level of 0.01 for autumn, whereas in summer, Tmin at the SSP2-4.5 midterm and both Tmax and Tmin under SSP5-8.5 showed increasing significant trends at the 0.01 level of significance. In spring, except for Tmax under SSP5-8.5 near-term periods, the results showed a significant increase in both Tmax and Tmin in near- and mid-term periods under both scenarios at a significant level of 0.0, indicating that future temperature change is unequivocal (Table 13). This finding also agrees with several reports concerning annual and seasonal maximum and minimum temperatures in different parts of Ethiopia, which revealed an increasing trend (Alemu and Dioha, 2020b; Bayable et al., 2021; Dalle et al., 2023; Shigute et al., 2024) in Addis Ababa, South Ethiopia, West Harerge Zone, and the Upper Ganale River Basin of Ethiopia.

The study area experienced significant temperature changes, with some stations experiencing larger increases. The mean maximum temperature ranged from 24.7°C to 30.2°C in the base year. Future projections suggest a mean maximum temperature between 26.1°C and 31°C, with the maximum expanding towards the River Ganale source. This was supported by the study of climate impact on rainfall and temperature of Dawa sub-basin, which found annual Tmin increases of 2.94, 3.45, 3.21, and 3.59°c and Yearly Tmax increased by 2.61, 2.83, 2.71, and 3.36°c for RCP4.5 and RCP8.5, respectively (Bulti & Abegaz, 2024).

The availability of water resources in the region will be significantly impacted by this temperature shift in the future because high temperatures lead to high evapotranspiration. This will worsen the river basin's hydrological cycle, ecosystem imbalance, and water scarcity. Temperature changes brought about by climate change are the consequence of both human and natural disruption of the natural equilibrium through excessive emissions of greenhouse gases. The availability of water resources in the river basin, as well as crop growth and biodiversity, will be severely impacted by high temperatures in both Tmax and Tmin.

1.1.2 Projection of Rainfall

The Upper Ganale River's highest rainfall occurred from 1985 to 2014, with the lowest around Chenemsa. Seasonal precipitation was highest in spring and autumn at Bore and Negele, with maximum and minimum values in May and March. Station-wise, historical data results indicated that the highest and lowest seasonal precipitation occur in spring and autumn, with 516.46 mm, 414.57 mm, 151 mm, and 111 mm at BOR and NGL, respectively. In terms of monthly rainfall, the spring maximum and minimum values are found in BOR and NGL, at 232.76 mm and 18 mm in May and March, respectively. The forecast for rainfall spatial dispersal in the UGRB shows variability in distribution compared to the past. Lower rainfall in the western and eastern parts is expected to decrease in the future, while the highest rainfall in the upper part of the basin is expected to increase. Moderate rainfall is expected to replace the lower rainfall areas. According to the spatially aggregated projection, annual rainfall showed a remarkable change relative to the base year. Thus, under SSP-5-8.5, the highest annual rainfall deviation (56%) resulted in the upper part of the basin. However, under SSP-5-8.5 in the outlet part of the UGRB, NGHL showed deviations of 51%, which differ from the results of other lowland stations. For more information, please refer to the previous publication ALI et al. (2025).

Table 14 presents divergent precipitation change from the six CMIP6 models under SSP2–4.5 and SSP5–8.5. The CanESM5 project results showed a significant decrease in the near-term by 6.07 but showed an increase in the mid-term under SSP2-4.5 by 15.4, SSP5-8.5, near-term, near-term, and mid-term results for CanESM5 models indicated a substantial increase of 5.49 and 45.82, respectively (Table 14).

Table 9. Annual Climate Change Comparison among the CMIP6 Models with respect to the base period

Scenario	SSP2-4.5		SSP5-8.5	
Timeframe	Near term	Midterm	Near term	Mid term
CanESM5	-6.07	15.42	5.49	45.82
FGOALS-g3	-8.09	-18.23	-5.70	-9.86
GISS-E2-1-G	-5.09	-20.94	-5.62	-5.03
INM-CM4-8	6.22	2.74	20.31	31.36
INM-CM5-0	-0.04	0.23	3.99	29.01
KACE-1-0-G	-7.77	-15.24	-7.91	-0.88
Ensembles	-3.47	-6.00	1.76	15.07

FGOALS-g3 and GISS-E2-1-G consistently show declines, with mid-term SSP2-4.5 reaching decreases of 18.23 and 20.94, respectively. The INM-CM4-8 and INM-CM5-0 models projected a positive change, especially under SSP5-8.5, with a mid-term increasing result of 31.36 and 29.01, respectively. On the other hand, KACE-1-0-G shows declines across most scenarios, though mid-term SSP5-8.5 approaches neutrality (-0.88). The overall ensemble mean was moderately negative under SSP2-4.5, with values decreasing to

3.47 near-term and 6.0 mid-term, but positive under SSP5-8.5, with increases of 1.76 near-term and 15.07 mid-term, as indicated in Table 14.

1.2 Climate Change Impacts on Water Resources

1.2.1 Historical Water Resource Availability

The water balance results for UGRB for each sub-basin were extracted from the QSWAT+ results for the base period. The historical water balance for the UGRB ensemble average summary is shown in Table 15, which includes precipitation, surface runoff, percolation (PEC), evapotranspiration (ET), lateral flow (Lat), water yield (WYLD), and potential evapotranspiration. In doing so, the percentage of the basin covered by precipitation was determined to assess water resource availability for irrigation developments. Accordingly, the base-year hydrological components showed that UGRB received 347.9 MCM of precipitation annually, with the highest rainfall in April (47 MCM) and the lowest in January (9.4 MCM). Similarly, surface runoff contributed 47.8 MCM annually, peaking between May and September (6–8 MCM) and falling during dry months. Water yield reached 56.7 MCM, closely following precipitation peaks, while lateral flow was comparatively small (9.0 MCM annually). At 115.8 MCM, percolation was the predominant subsurface component, maintaining groundwater recharge. About half of the precipitation, or 175.6 MCM, was used by evapotranspiration. There was a noticeable water deficit because potential evapotranspiration (PET) was significantly higher (336.1 MCM annually). Only about 82% of PET was covered by PCP, with significant deficiencies during dry months (January 23 percent, December 32 percent) (Table 15).

Table 10.Estimation of Water Balance of Upper Ganale River sub-basin (from 1988 to 2014)

Water (MMC)	Balance	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
PCP		9.4	13.0	25.0	46.5	47.0	30.6	28.2	37.8	39.3	40.9	19.2	10.9	347.9
Surq.gen		0.3	0.3	0.8	2.6	6.3	5.7	4.2	8.2	8.3	7.9	2.5	0.7	47.8
PERC		0.4	0.2	0.2	0.4	0.9	1.1	0.9	0.9	1.1	1.2	1.1	0.6	9.0
WYLD		0.6	0.6	0.9	3.0	7.2	6.8	5.0	9.1	9.4	9.1	3.6	1.4	56.7
Latq		4.3	2.3	2.0	5.2	12.6	14.0	11.4	12.0	13.7	16.1	14.0	8.3	115.8
ET		17.7	17.2	16.0	15.1	14.7	12.8	11.3	11.9	13.4	12.8	15.0	17.6	175.6
PET		40.4	42.5	35.0	25.9	24.7	22.6	19.8	20.6	23.0	21.3	25.9	34.3	336.1

1.2.2 Climate Change Impacts on Water Resources

The simulated hydrological responses for the base period (1985-2014) and the future periods (near-term and mid-term) were then compared to evaluate changes in the watershed's future water balance components. The results of the impact of future climate change on water balance components, based on simulated precipitation, total water yield, surface runoff, streamflow, evapotranspiration (ET), groundwater flow, and lateral flow from the QSWAT+ model, are presented in Table 16. The results indicated a reduction in precipitation in the near term and mid-term for SSP2-4.5 scenarios by (-3.478%) and (-6.087%), respectively. In SSP5-8.5 scenarios, both in the near term and the mid-term, there is an increase in rainfall, as the ensemble model results project by 1.786% and 15.51%, respectively. Regarding the annual surface runoff percentage changes under SSP2-4.5 near- and mid-term, and SSP5-8.5 near-term, the results showed reductions of (-1.274%), (-24.848), and (-4.459), respectively. During the SSP5-8.5 mid-term period, surface runoff increases by 21.65% (Table 16). The Evapotranspiration results show that the future is warming, with the percentage change in ET showing a steady incremental trend.

Table 11.The Annual Projected Water Balance and Percentage Change from the Base Period

Scenarios		*SSP2-4.5SSP245*		*SSP5-8.5SSP585*	
Period	Base period	Near Term	Mid Term	Near Term	Mid Term
Rainfall(mm)	1150	1110	1080	1170	1340
% Change		-3.478	-6.087	1.786	15.517
Surq_gen(mm)	159	155	118	150	191
% Change		-1.274	-24.848	-4.459	21.656
Latq(mm)	18.3	15.4	13.1	15.7	18.1
% Change		-0.25	-0.448	-0.224	-0.017
Wateryld(mm)	187	170	131	166	209
% Change		-1.734	-24.277	-4.046	20.809
perc(mm)	386	225	205	252	287
% Change		-23.469	-30.272	-14.286	-2.381
et(mm)	583	713	743	748	842
%Change		2.59	6.906	7.626	21.151

Climate change projections have significant implications for water management in the Upper Genale Basin. According to the result of Table 16, under the SSP2-4.5 pathway, both the near and mid-term periods showed a reduction in rainfall. This result is in line with the finding of Legass et al. (2026) that stated reduced rainfall in parts of Ethiopia under moderate emission scenarios. As indicated in Table 16 results, the reduced rainfall drives declines in runoff, water yield, and percolation, pointing to heightened drought risk and reduced water security. Conversely, SSP5-8.5 mid-term projects a more rainfall result with rainfall increases exceeding 15%. This aligns with recent CMIP6-driven SWAT+ simulations in the Awash River Basin, which report intensified wet season rainfall under high-emission scenarios and other findings in East Africa(Assamnew and Tsidu, 2025; Bedada et al., 2025).

From the result of the study, both SSP pathways show stronger hydrological responses in the mid-term compared to the near term. SSP5 8.5 mid-term is the most extreme, with precipitation and runoff values far exceeding baseline, pointing to greater flood risk and soil erosion potential. The seasonal cycle remains intact (wet peaks in April–May and September–October), but the magnitude of rainfall and runoff is amplified. This means existing water management strategies must adapt to more intense wet seasons. Increased percolation and water yield under mid-term scenarios suggest improved recharge opportunities. However, the concentration of recharge in fewer months could stress aquifers during prolonged dry periods. The overall study results imply that flood control infrastructure will be critical under SSP5 8.5 mid-term conditions. On the other hand, the intensification of the impact of climate change due to increasing temperature and reduction of rainfall in the mid-term significantly affects the water resources of the upper Genale River basin. Water stress, especially due to the reduction of total water yield, runoff, and groundwater flow during the rainy season, may cause extensive impacts on water supply, agricultural activities, and ecosystem services.

1.3 Climate Change Effect on Streamflow

The trained hybrid and individual ML models were utilized to project future streamflow under CMIP6 climate change scenarios (SSP245 and SSP585) for Near-term and Mid-century periods. An iterative bootstrapping mechanism was implemented for forecasting, where the model's prediction for day t was fed as a lagged input for day t+1, enabling continuous forecasting through 2074. Placeholder historical mean values were used for QSWATPLUS simulated flow in the future projections. From the result, future predictions detail summary results were illustrated in Table 17. The result indicated that the QSWATPLUS predicted streamflows showed a decreasing trend in the near term under both SSP245 and SSP585, with 7.73 and 35.81 %, respectively. This result is compatible with the finding of the study by Shigute et al. (2024) in UGRB, which stated that the projected increase and decrease in streamflow during the dry and wet seasons under RCP4.5 and RCP8.5 are attributed to the likely increase and decrease in rainfall amounts during these respective periods.

Whereas, an increasing aspect under mid-term for both SSP245 and SSP585 of about 10.19% and 20.9%, respectively. With respect to ML results, the sample illustration in Table 17 results by TSLM, QSWATPLUS-TSLM, Stacked ensemble, and BMA by scenario and period reveals varying magnitudes and distributions of projected flow, with general increases expected under both SSPs, except under near-term SSP245, for all models. As a result, QSWATPLUS-LSTM, Stacked-Ensemble, BMA, and LSTM showed decreasing aspects of -12.18%, -4.38%, -7.7%, and – 9.8%, respectively. This indicates ML and QSWATPLUS predictions coincide with the increase of mid-term periods, and in near-term periods, BMA coincides in decreasing aspects with QSWATPLUS.

Table 12. Predicted percentage changes in annual Streamflow under near and midterm for selected ML, Stacked, BMA, and QSWATPLUS predicted results in m^3/s

Period	Scenario	LSTM	QSWATPLUS-LSTM	Stacked Ensemble	BMA	QSWATPLUS	LSTM% Change	QSWAT-LSTM% Change	Stacked Ensemble% Change	BMA % Change	QSWATPLUS % Change
Base period		134.6	134.6	134.6	134.6	134.6					
Midterm	ssp245	146.28	142.59	157.5	148.5	148.3	8.68	5.94	16.99	10.30	10.19
	ssp585	184.	179.6	193.8	178.4	162.7	36.71	33.44	43.99	32.54	20.90
Near-term	ssp245	121.4	118.2	128.7	124.2	124.2	-9.80	-12.18	-4.38	-7.70	-7.73
	ssp585	142.0	138.4	147.6	140.1	86.40	5.50	2.83	9.66	4.05	-35.81