Journal of
Water & Climate Change

IWA PUBLISHING

Journal of Water and Climate Change Vol 14 No 5, 1723 doi: 10.2166/wcc.2023.038

Study of hydrologically critical subbasins under climate change

Shishir Gaur[a], Ranveer Kumar[a,*], Anurag Ohri[a], Shreyansh Mishra[a], Ajeet Kumar Gond[a], Shyam Bihari Dwivedi[a], Medha Jha[a], Abhyanand Chaturvedi[b] and Bhola Nath Singh[a]

[a] Department of Civil Engineering, Indian Institute of Technology (BHU), Varanasi, India
[b] Hindalco Industries Limited (Mahan Aluminium), Bargawan, Madhya Pradesh, India
*Corresponding author. E-mail: ranveerkumar.rs.civ20@iitbhu.ac.in

RK, 0000-0001-7530-3581

ABSTRACT

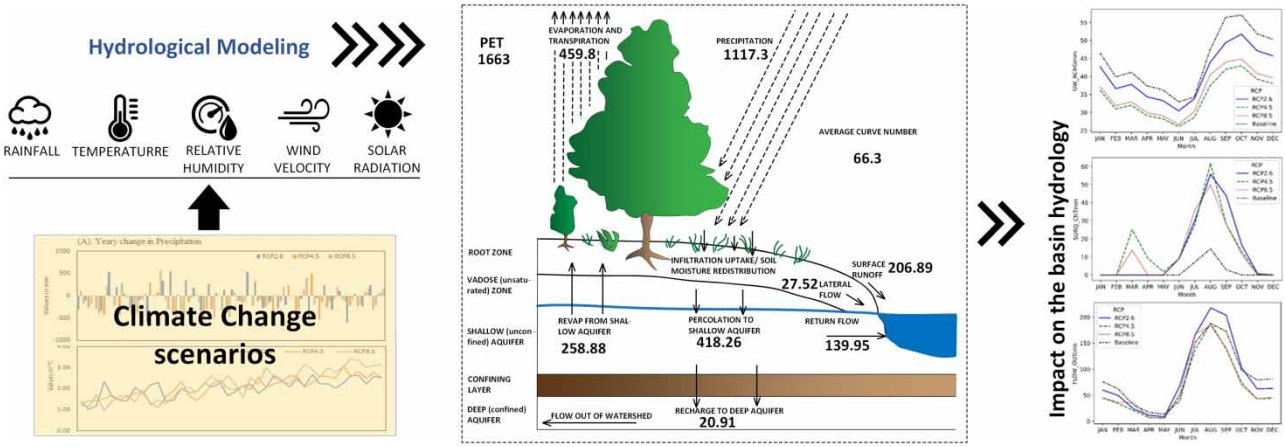
This study aims to evaluate the impact of climate change on the surface water hydrology of the Gopad river basin in India. The outputs of four CMIP6 Global Climate Models have been downscaled using the statistical downscaling method to the basin level. A comparative analysis for the accuracy achieved in the bias correction for the combination of GCM and downscaling method has been performed before utilising the downscaled weather parameters for hydrological study. The MIROC6 and ACCESS-CM2 were found best for the simulation of precipitation and temperature, respectively. The Distribution Mapping and Variance Scaling methods have shown better accuracy w.r.t other statistical methods. The impact of climate change has been found significant since the temperature has been observed to be increased by 3.16 °C by the end of 2060; meanwhile, there is an average decrease of 9.2% in the annual rainfall from the baseline. The peak runoff has increased while there is a significant decrease in the groundwater recharge. Further, hydrologically critical subbasins (HCS) have been delineated based on the runoff, groundwater recharge, and baseflow. Most HCS was observed to be situated in the upper Gopad river basin, representing the area's pristine conditions.

Key words: bias correction, climate change impact, GCM, hydrologically critical subbasins, statistical downscaling

HIGHLIGHTS

- The ACCESS-CM2 and MIROC6 are more accurate for the study area.
- The Variance Scaling and Distribution Mapping methods have shown better accuracy than others.
- A rise in average temperature has been observed, while there is a decrease in annual rainfall from the baseline.
- The area with less human intervention has proven to be hydrologically critical for the Gopad river basin.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Water availability considerably influences the ecology, power generation and human activities within a river basin. The stored water in the aquifer feeds the river as base flow and percolates deeper as deep aquifer recharge. Precipitation and temperature are the most influential variables in a hydro-environmental system. As a result, monitoring their changes in the future might assist decision-makers in dealing with issues such as drought, flash floods, and high evapotranspiration. Most parts of the world have documented the effects of climate change on hydrologic processes such as streamflow, baseflow, surface runoff, and evapotranspiration. According to Richmond and Yohe (Melillo *et al.* 2014), extreme precipitation occurrences in the Midwest USA have increased over the previous century, resulting in increased surface runoff and streamflow. Furthermore, as climate forecasts show, the frequency and amplitude of such extremes may continue to climb in the coming decades (Sun *et al.* 2016).

Hydrologically Critical Subbasins (HCS) are areas significantly contributing to hydrological variables. Several watershed hydrogeological parameters, such as soil, slope, soil moisture content, topography, rainfall intensity, rainfall depth and frequency, all are transient and significantly influence the characteristics of subbasins (Needelman *et al.* 2004). Field methods (Mehta *et al.* 2004) and surface water modelling approaches (Niraula *et al.* 2013) have been utilised to find the critical areas for runoff and nutrients. Apart from runoff, the groundwater recharge and baseflow contribution of a subbasin are significant factors for water resource management, and these parameters should also be analysed. The low groundwater recharge areas are better for runoff generation but contribute less to groundwater storage. The majority of groundwater flow is replenished by the areas of high groundwater recharge which are limited. These pristine areas with high percolation rates are critical for the basin.

The sixth phase of Coupled Model Intercomparison Projects (CMIP) is now available with daily simulation output of 58 General Circulation Models (GCMs) across the globe. Regarding the number of modelling organisations' participation, the number of future scenarios considered, and the number of different experiments undertaken, CMIP6 (Eyring *et al.* 2016) marks a significant increase over CMIP5. CMIP6 future climate forecasts include enhanced emissions, improved model parameterisation, land use scenarios, and physical processes, among other things, all driven by scenarios based on shared socioeconomic paths (SSPs) (Eyring *et al.* 2016). In the simulation of distinct climate variables in different regions, CMIP6 GCMs have shown both better and worse performance than CMIP5. The performance of CMIP6 and CMIP5 in simulating the Indian summer monsoon rainfall was inconsistent in a comparative study by Gusain *et al.* (2020). However, the performances of CMIP6 models in the simulation of global temperature extremes, diminishing precipitation and droughts have been found to be superior (Rivera & Arnould 2020).

The uncertainties in the GCMs from different sources, including mathematical formulation, assumptions, model resolution and calibration technique, limit the use of all available GCMs for climate projection in a different region (Sun *et al.* 2018). Therefore, the models must be compared for their ability to project local climate scenarios. The statistical downscaling of GCMs output is the most inexpensive method to cater for the demand of ongoing climate impact studies. However, these methods are frequently tainted with limitations and errors due to assumptions and approximations associated with each method (Maraun *et al.* 2019). A comparative assessment of the relative accuracy in the downscaling bias correction is required before utilising the GCMs data for hydrological modelling.

Forecasts from climate models have been coupled with hydrologic models to estimate the potential consequences of climate change on water systems (Mohammed *et al.* 2015). Recently, Muto *et al.* (2022) used an ensembled climate dataset to the SWAT (Soil & Water Assessment Tool) model to quantify the climate change impact on the hydrology of the Tokoro River Basin. Kumar *et al.* (2022) studied the climate impact on the upper Betwa River catchment with the SWAT model and found a decrease in the annual rainfall and surface runoff for all scenarios. In India, the SWAT model has been used extensively with downscaled GCM data to study the impact of climate change on the hydrology of a basin (Sharannya *et al.* 2018). All these studies show a notable increase in rainfall intensity in the monsoon period and an increase of 1 °C–3 °C in the air temperature.

The novel methodology has been implemented in the Gopad River basin to delineate the critical source areas (critical subbasins) for streamflow under the impact of climate change. The accuracy of CMIP6 models, along with the statistical downscaling methods, has also been compared. Four GCMs have been selected based on the performance study done by Kamruzzaman *et al.* (2021) (ACCESS-CM2 (Australian Community Climate and Earth-System Simulator), INM-CM5-0 (Institute of Numerical Mathematics, Russia), ACCESS-ESM1-5 (Australian Community Climate and Earth-System Simulator),

and MIROC6 (Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan)). The climate model outputs have been downscaled by statistical methods, and the parameters have been used as weather parameters input in a calibrated SWAT model. The hydrological simulation outputs have been analysed to find the hydrologically critical areas (subbasins) for runoff, recharge and baseflow contribution.

2. STUDY AREA

Gopad river is located in the heart of the Sone River basin, encapsulating an area of 6,036.3 km^2 (Latitude 24° 40′ to 23° 20′; Longitude 81° 40′ to 82° 40′) (Figure 1). The area is more or less pristine compared to other river basins in the Sone River watershed. The site is sparsely populated, and most of the area is forest and grassland (64.43%). Also, about 10% of the

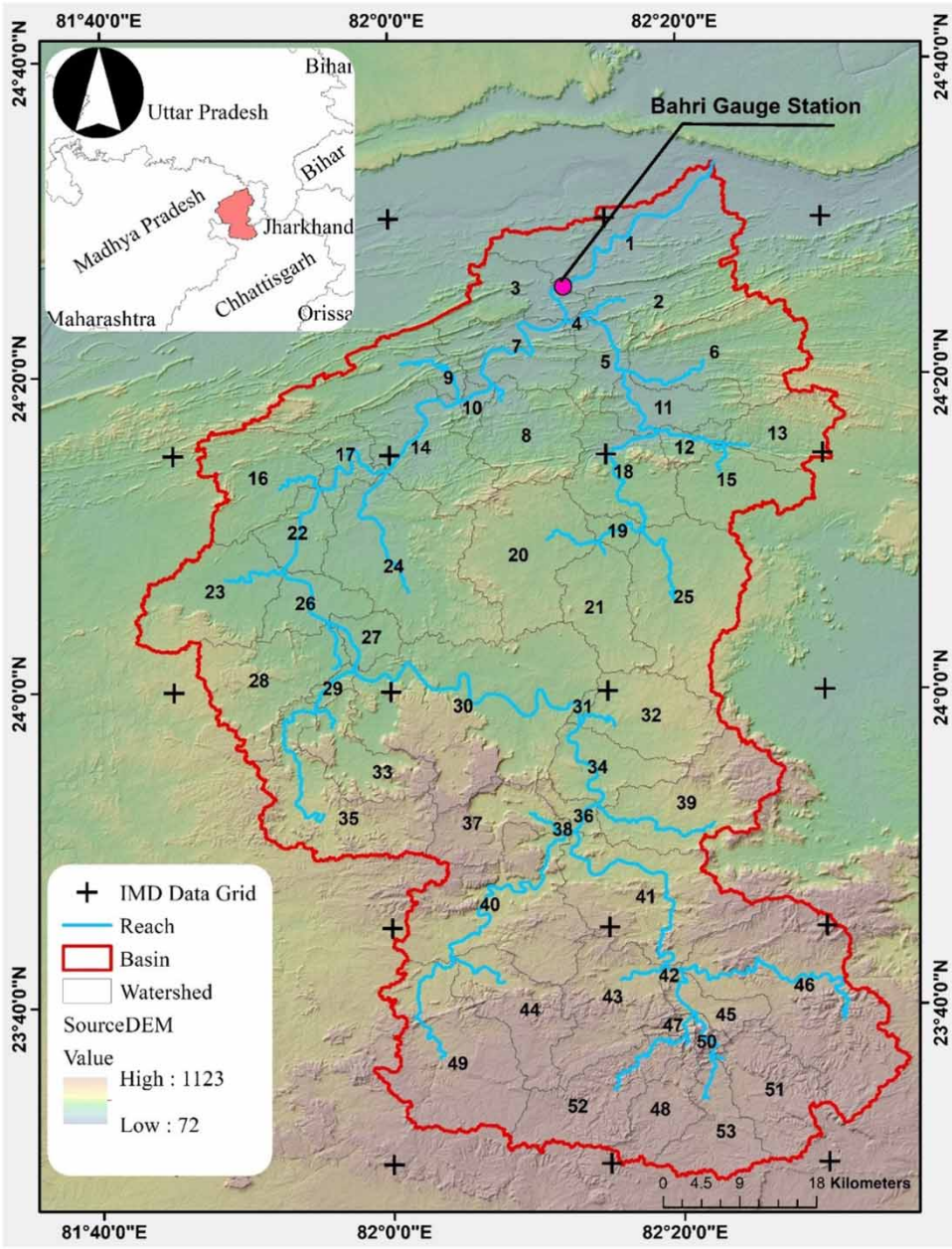


Figure 1 | Study area.

area is barren or sparsely vegetated with seasonal grass. Agricultural activities are limited to the lower part of the basin (23.5% of the total basin area). It is a perennial river which receives its lean period discharge from the adjoining aquifers as base flow. There has not been any study on hydrology and the impact of climate change.

The precipitation data from the Indian meteorological department (IMD) (Pai *et al.* 2014) shows a mixed trend of yearly precipitation with a mean yearly rainfall of 1,117 mm from 1980 to 2020. The rainfall has shown a variation of a minimum of 734 mm (2017) to a maximum of 1,753 mm (1994). Maximum rainfall occurs during the southwest monsoon season. The monsoon season, which lasts from June to September, accounts for 89% of the annual rainfall. The month of July is the wettest of the year. Only 11% of the yearly rainfall occurs between October and May. The temperature begins to rise in early February and reaches its peak in May. In June, the daily mean maximum temperature is 43 °C. The monsoon usually arrives around the middle of June, bringing a significant temperature reduction. In general, January is the coldest month of the year. The humidity is lowest in April, around 35% during the southwest monsoon; the humidity is higher due to heavy rains, reaching a maximum of about 85% in August. Because of the high temperatures, the humidity drops again in October, and the monsoon retreats. The daily mean annual relative humidity of the area is 66%.

3. METHODOLOGY

3.1. Climate dataset

GCMs are among the most advanced techniques, using transient climate simulations to simulate climatic conditions on Earth for the next hundreds of years. The GCMs under CMIP6 (O'Neill *et al.* 2016) for this study have been selected based on the comparison performed by Kamruzzaman *et al.* (2021). Four GCMs (ACCESS-CM2, INM-CM5-0, ACCESS-ESM1-5, and MIROC6) have been downscaled to IMD 2D grid resolution ($0.25° \times 0.25°$) with bilinear interpolation. The bias correction has been performed by statistical methods, using IMD 2D gridded dataset for three SSPs (SSP1 based on RCP-2.6, SSP2 based on RCP-4.5 and SSP5 based on RCP-8.5). The IMD has developed the gridded data for the research community's needs by interpolating the gaged data by inverse distance weighted method.

3.2. Statistical downscaling

The downscaling of the simulation outputs of GCM has been performed using five widely used statistical methods (1. Linear scaling (LS), 2. Local intensity scaling of precipitation (LI), 3. Power transformation (PT), 4. Variance scaling (VS) and 5. Distribution mapping (DM)) (Figure 2). The foundation for linear scaling bias correction is the perfect agreement between the monthly mean of observed and simulated values. Precipitation is adjusted using the ratio of long-term monthly mean observed and historical run data. In contrast, the temperature is adjusted using an additive term based on the difference between the long-term monthly mean observed and the historical run. Because linear scaling only compensates for monthly mean bias, it cannot correct wet-day frequency and intensity biases.

The local intensity scaling is a stepwise method to correct the biases due to the mean as well as both wet-dry frequencies and intensities of precipitation time series. In this study, the threshold precipitation in the LI method varied from 0.05 to 1.15 mm for different GCMs. The larger variation indicates a large number of low precipitations simulated events. The scaling factor is much higher for ACCESS-CM2 and ACCESS-ESM1-5, varying from 0.19 to 40.2 (higher values are more dominant from May to September). The scaling factor for MIRO-C6 and INM-CM5-0 varies from 0.29 to 4.87 (higher values are more dominant from August to October). The factor is more consistent for MIRO-C6, varying around 0.38–1.24. This represents that MIRO-C6 is much more consistent in estimating precipitation than other GCMs, as those overestimate the rainfall intensity.

Power transformation utilises mean as well as variance for the correction of discrepancies. As a result, a non-linear correction with an exponential form ($a.P^b$) is used to precisely modify the variance statistics of a precipitation time series (Leander *et al.* 2008). A 90-day frame centred on the interval is used to estimate parameter b on a monthly basis. The coefficient of variation (CV) of the corrected daily simulated precipitation (P^b) and the CV of the observed daily precipitation (P_{obs}) for each month are compared to estimate parameter b.

Chen *et al.* (2011) described a similar method to Power Transformation known as 'variance scaling,' which can be used to incrementally adjust the mean and variance of temperature time series. First, the mean of the simulated time series is adjusted using linear scaling. The mean-corrected historical and future scenario runs are then adjusted to a zero mean on a monthly basis. Distribution mapping aims to match the observed distribution function to the distribution function of RCM-simulated

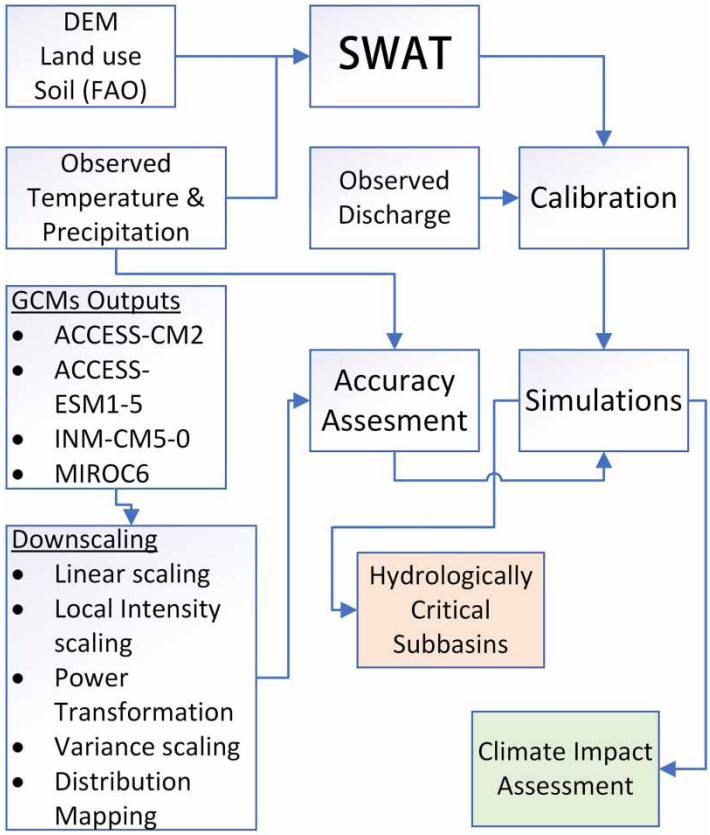


Figure 2 | Flow chart of the research work.

climatic variables. This may be accomplished by developing a transfer function that shifts the precipitation and temperature occurrence distributions. The downscaled average temperature and precipitation have been presented in Figure 3.

3.3. Evaluation statistics for statistical downscaling and SWAT model calibration

Multiple accuracy metrices provide a complete picture of the model performance since each metric provides different insights. A total of four accuracy metrices (RMSE, PBIAS, NS and R^2) have been used for the evaluation of SWAT model performance and the accuracy assessment of downscaling methods (Figure 4) to reduce the bias associated with each metric.

3.3.1. RMSE

Root mean squared error (RMSE) is an established way to quantify the error of a predictive numeric model. Simply, it calculates the deviation (spread) of error between the actual and forecasted datasets. Mathematically it can be represented as:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Q_{i,\,obs} - Q_{i,sim})^2}{n}} \tag{1}$$

where $Q_{i,sim}$ are the simulated values (discharge or climate parameter); $Q_{i,\,obs}$ is observed values of the parameter (discharge or climate parameter), and n is the number of observations available for analysis. It is widely applied to compute the prediction error since it is: (1) A single number to judge from the training to the deployment phase of a model. (2) Representing the relative size of the error from each iteration to the next and not just the absolute magnitude of the error.

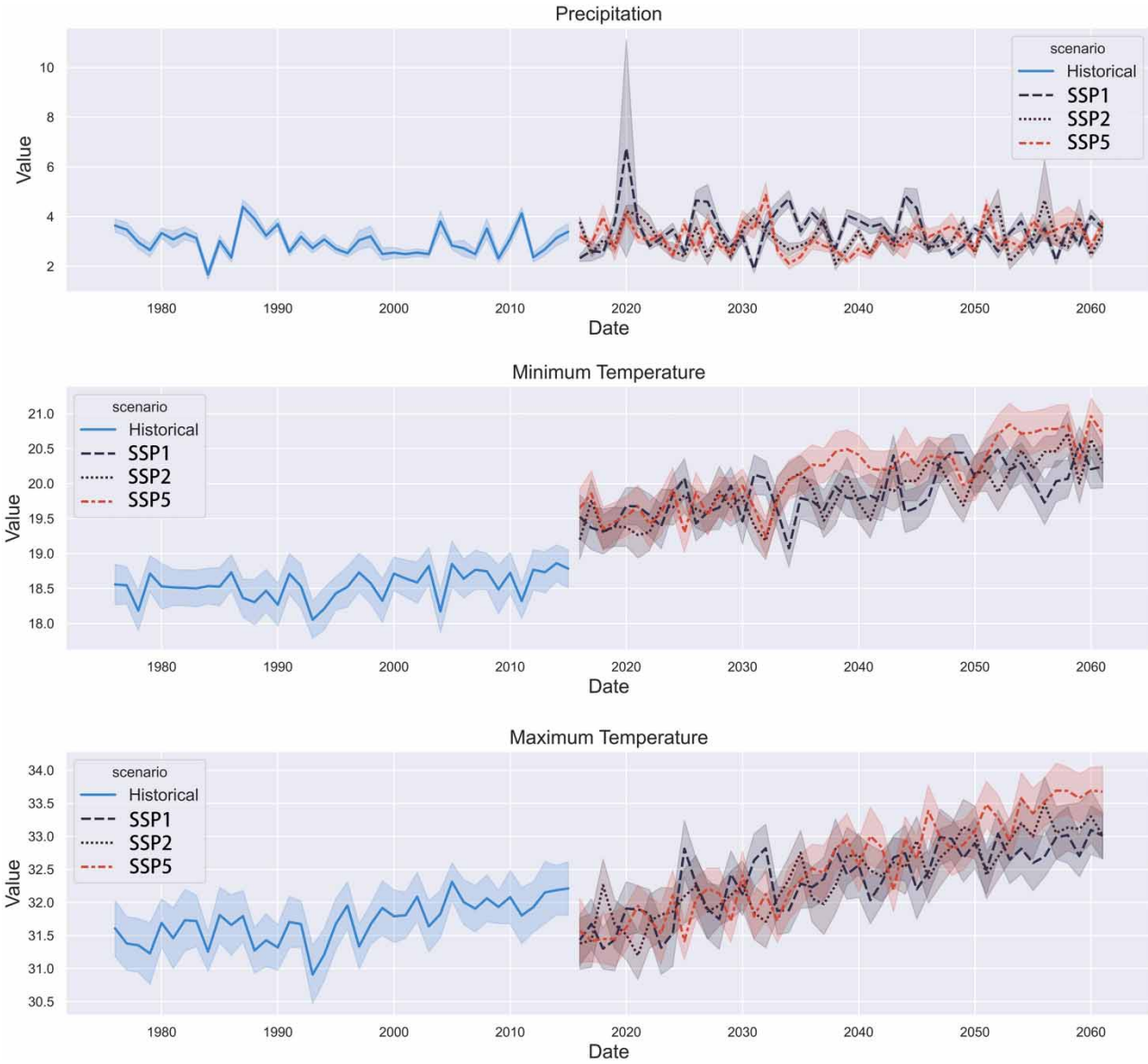


Figure 3 | Variation of precipitation, minimum temperature and maximum temperature for IPCC scenarios (SSP1, SSP2, and SSP5) along with the historical data.

3.3.2. PBIAS

The average tendency for the simulated data to be higher or smaller than the observations is represented by per cent bias (PBIAS). Zero is the best value, whereas smaller magnitude values suggest better simulations. Positive numbers suggest that the model is underestimated, whereas negative values indicate that the model is overestimated.

$$PBIAS = 100 \times \frac{\sum_{i=1}^{n} (Q_{obs} - Q_{sim})_i}{\sum_{i=1}^{n} Q_{sim,i}} \quad (2)$$

where Q is a variable (discharge, temperature or precipitation), '*obs*' and '*sim*' represent observed and simulated data, respectively.

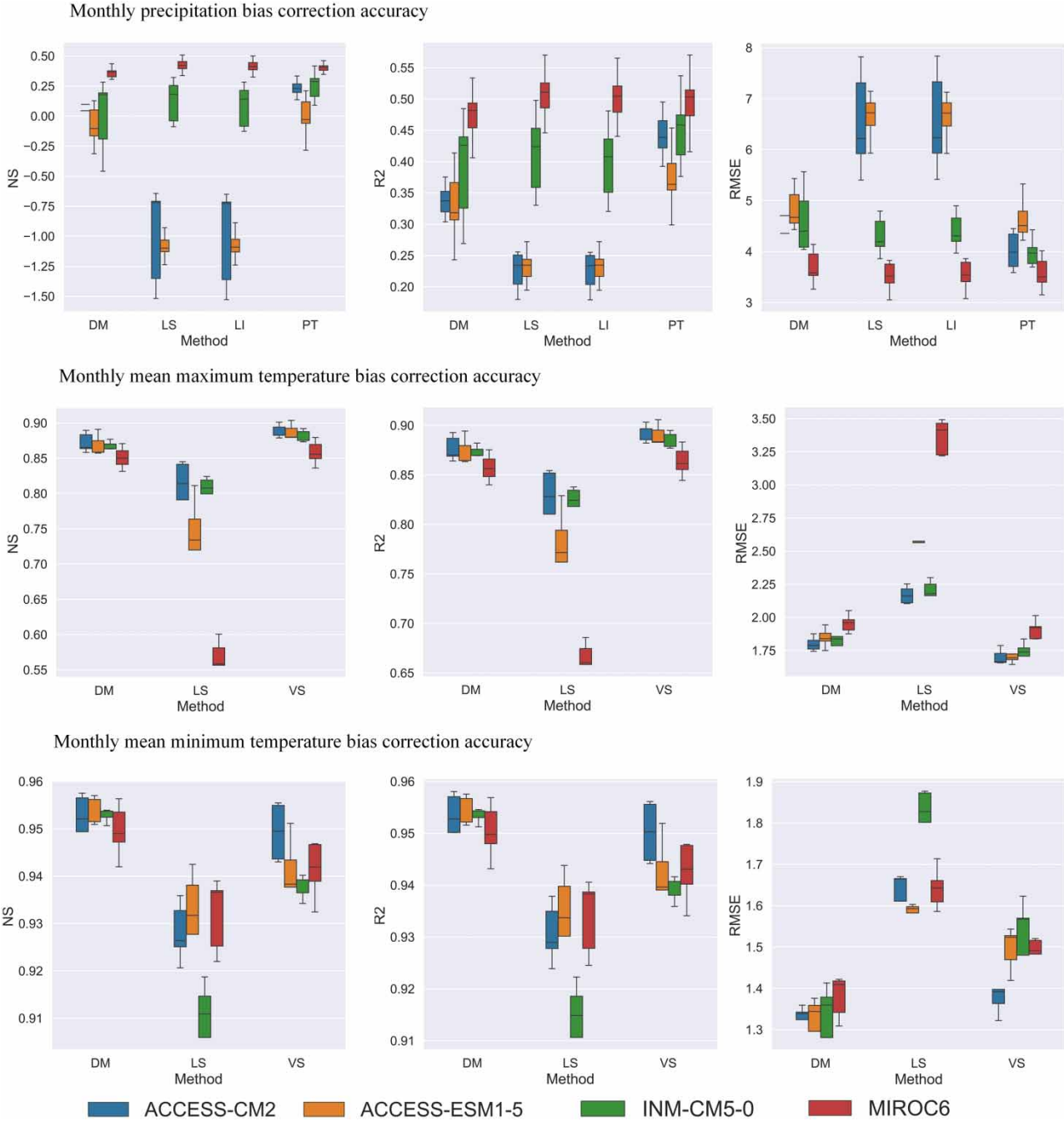


Figure 4 | Accuracy variation among all the grid points in the Gopad river basin.

3.3.3. Nash–Sutcliffe efficiency (NS)

One of the reliable measures to assess the predictive skill of hydrological models, NSE normalises model performance into an interpretable scale. It is described as:

$$NS = 1 - \frac{\sum_{i=1}^{n} (Q_{obs} - Q_{sim})_i^2}{\sum_{i=1}^{n} (Q_{obs,i} - \bar{Q}_{obs})_i^2} \tag{3}$$

where Q is a variable (discharge, temperature or precipitation), '*obs*' and '*sim*' stand for observed and simulated data, respectively. $\bar{Q}_{obs}$ is the mean of the observed variable. The numerator of Eq (3.29) is the sum of the squares of the residuals. In contrast, the denominator is the sum of squares of the simulated values about the mean of the simulated values, representing the overall variation of the observed values about the mean (worst value $= -\infty$; best value $= 1$).

3.3.4. R^2

The coefficient of determination (R^2) is a statistical measure for linear regression models that indicates the proportion of variance in the response variable(s), which is illustrated by the independent variable(s). In other words, it measures the combined dispersion against solo dispersion of the observed and simulated series and the wellness of fit of a regression model. Usually, the higher the R^2 value, the better the regression model fits observations.

$$R^2 = \frac{\left[\sum_{i=1}^{n}(Q_{obs,i} - \bar{Q}_{obs})(Q_{sim,i} - \bar{Q}_{sim})\right]^2}{\sum_{i=1}^{n}(Q_{obs,i} - \bar{Q}_{obs})^2 \sum_{i=1}^{n}(Q_{sim,i} - \bar{Q}_{sim})^2} \tag{4}$$

where Q_{sim} is the simulated i^{th} value, and the Q_{obs} element is the observed i^{th} value (worst value $= 0\%$; best value $= 100\%$).

3.4. Hydrological modelling and calibration

3.4.1. SWAT model

SWAT (Arnold *et al.* 2012) is a physically based, computationally efficient watershed-to-basin scale surface water long-term yield model developed by Dr Jeff Arnold for Agricultural Research Service, USA. The SWAT (Soil and Water Assessment Tool) model is a widely-used hydrological model for assessing the impact of land use and management practices on water resources. The stream flow for mid-range, dry and low flows is better simulated by SWAT as compared to other models (Tegegne *et al.* 2017).

The basin is divided into several sub-basins. All input information is grouped into five categories: climate; hydrologic response units (HRUs); ponds/wetlands; groundwater, and main channel or reach, draining the subbasin. Within the subbasin, land regions are grouped into HRUs with certain combinations of soil, land cover, and management. SWAT simulates the watershed hydrology based on climate, water quality, soil and land cover parameters in two major divisions. The first is the land phase of the hydrologic cycle, which is based on the water balance equation:

$$SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{gw}) \tag{5}$$

where SW_t is the final soil water content (mm H_2O), SW_0 is the initial soil water content on the day i (mm H_2O), t is the time (days), R_{day} is the amount of precipitation on the day i (mm H_2O), Q_{surf} is the amount of surface runoff on the day i (mm H_2O), E_a is the amount of evapotranspiration on the day i (mm H_2O), w_{seep} is the amount of water entering the vadose zone from the soil profile on the day i (mm H_2O), and Q_{gw} is the amount of return flow on the day i (mm H_2O).

The second phase is the routing phase of the hydrologic cycle. SWAT estimates the loadings of water, sediment, nutrients, and pesticides to the main channel, which are then channelled through the watershed's stream network. SWAT simulates the change of chemicals in the stream and streambed in addition to tracking mass movement in the channel.

In the present study, the Gopad river basin is delineated to Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM). For flatter terrain in the lower part of the basin, the main stream is manually digitised and burned to the DEM to remove errors due to flat topography in stream delineation. The watershed was delineated with a threshold flow accumulation of 1,000 hectares of draining area. A total of 53 subbasins and stream reaches, along with the subbasin parameters (length, width, area, elevation), have been calculated. The land use data has been prepared with Landsat 8 imagery using maximum likelihood classification. The soil data is imported from Harmonized World Soil Database (HWSD) by the Food and Agriculture Organization (FAO) of the United States. The land cover and soil data have been used along with four slope classes (<5%, 5–10%, 10–15%, >15%) to calculate the HRUs based on dominance criteria. Climate data has

been imported from Indian Meteorological Department (IMD) 2D gridded data and converted to SWAT input txt file. The model is simulated for 1975–2006 with five year warmup period (Figure 5).

3.4.2. Calibration

The SWAT Model has been calibrated with observed monthly mean river discharge using Sequential Uncertainty Fitting (SUFI-2) in SWAT-CUP (Abbaspour 2019). SWAT-CUP provides an interactive graphical user interface to calibrate SWAT model parameters with multiple observation data. The Nash-Sutcliffe objective function has been optimised with 4,000 iterations. 16 parameters with five soil properties of the different layers have been optimised. Among the optimised parameters, CN2 (Curve Number) has been the most sensitive (Figure 6), followed by GWQMN (threshold depth of water in the shallow aquifer required for return flow to occur) and GW_REVAP (Groundwater revap coefficient). The optimised parameters ranked according to their sensitivity are represented in Table 1. The mean monthly discharge data at Bahri gauge station (subbasin 3) from 1980 to 2007 has been used for calibration. The discharge at the given gauge site has shown mixed variation over the measured time period, in accordance with the rainfall in the area, with a peak discharge as high as 1,400 m^3/sec and as low as 2 m^3/sec.

3.5. Simulation with future climate data

The GCMs outputs have been downscaled with the help of the CMhyd tool. The IMD 2D gridded data has been utilised as observed data for bias correction. All the downscaled data has been analysed for accuracy compared to the IMD data. The most accurate method and GCM model combination are used to estimate the future IPCC scenarios (SSP1, SSP2 and SSP5). The downscaled time-series precipitation and temperature have been used to simulate the calibrated SWAT model. In this research, the downscaled precipitation from MIROC6 GCM with the help of power transformation and temperature from ACCESS-CM2 downscaled with distribution mapping has been used for SWAT simulations.

3.6. Hydrologically critical subbasins

The source of dry weather flow in the Gopad River is majorly satisfied by the base flow. The subbasins contributing significantly to the baseflow are critical and need to be preserved for the flow of the Gopad river. The runoff, recharge, and baseflow contribution from each subbasin has been analysed based on runoff or recharge generated per unit of rainfall. Threshold

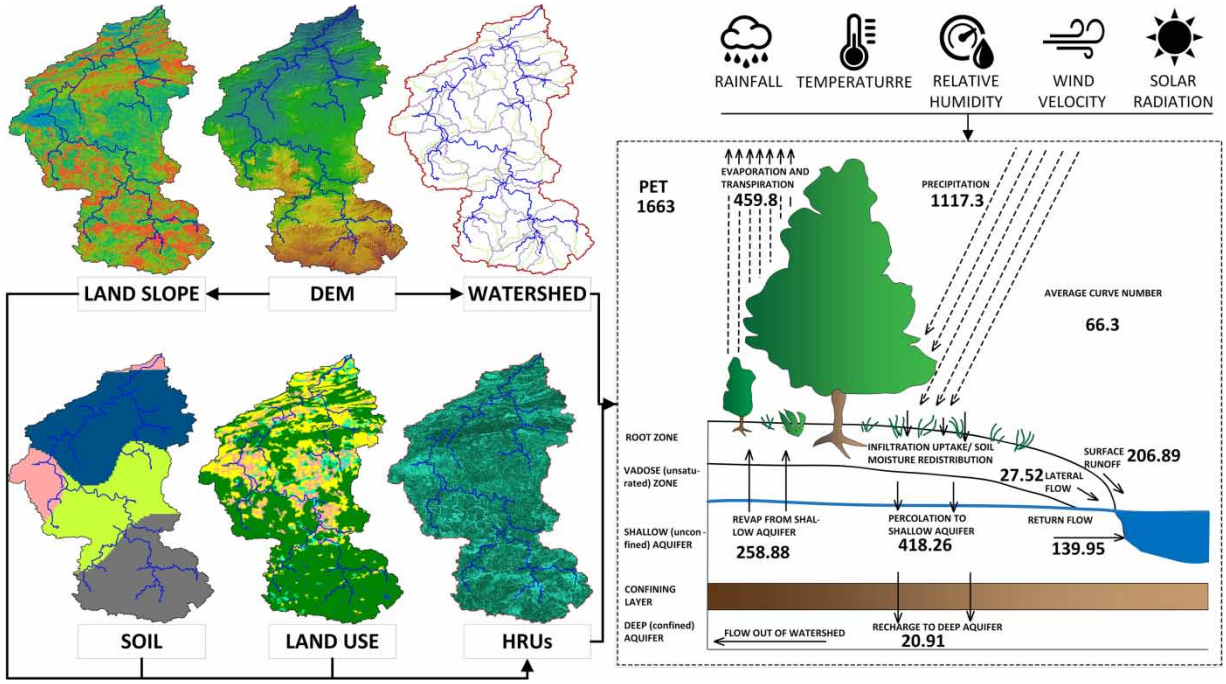


Figure 5 | SWAT Model of Gopad river basin; the value represents the average basin surface water budgets for the simulation period (1975–2006).

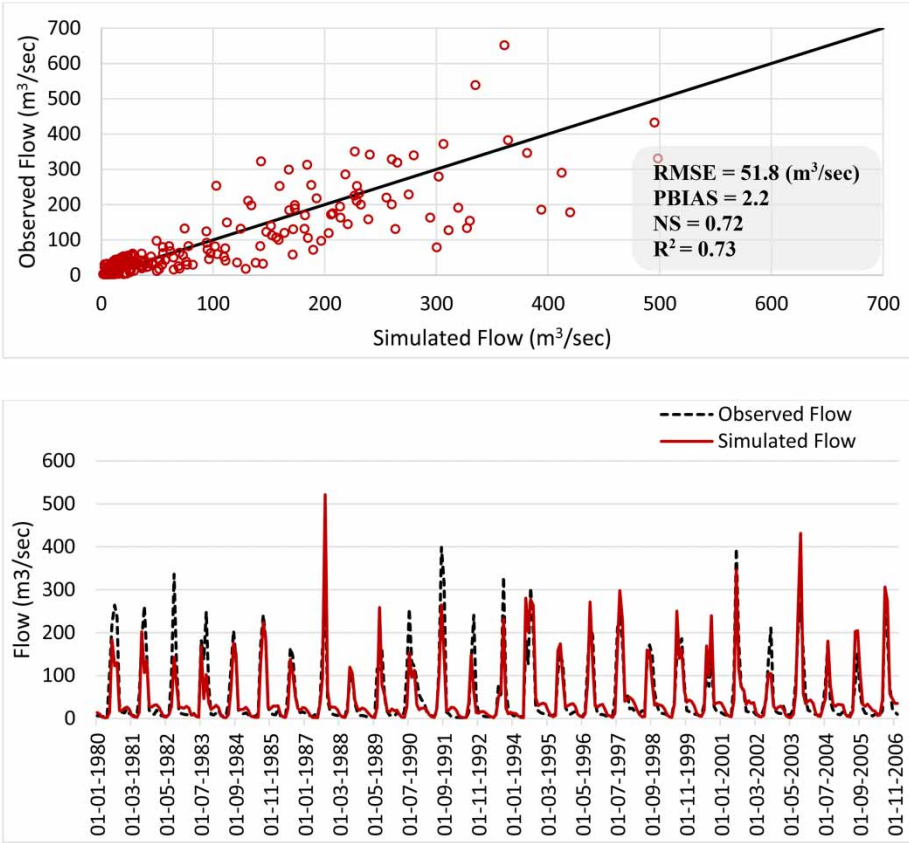


Figure 6 | SWAT calibration summary statistics and observed vs simulated discharge at Bahri gauge station.

Table 1 | Parameters after SWAT model calibration

Rank	Parameter	Minimum	Maximum	Fitted Value
1	CN2.mgt	− 0.255	0.048	− 0.143
2	GWQMN.gw	0.00	2,734.664	510.698
3	GW_REVAP.gw	0.099	0.2	0.167
4	GW_DELAY.gw	217.935	500.00	483.146
5	REVAPMN.gw	0.00	287.447	149.401
6	CH_K2.rte	0.01	95.717	9.126
7	ESCO.hru	0.589	0.796	0.605
8	OV_N.hru	0.082	0.447	0.286
9	SLSUBBSN.hru	− 0.791	0.069	− 0.776
10	ALPHA_BF.gw	0.284	0.453	0.395
11	SURLAG.bsn	8.536	24.00	21.382

values of 0.082, 0.402 and 0.488 have been used to identify the subbasins with runoff/rainfall, groundwater recharge/rainfall and baseflow/streamflow ratios as the critical subbasins. The threshold corresponds to the 50th percentile value of the respective ratios. The subbasins critical for all three ratios were identified as hydrologically critical subbasins for the study area.

4. RESULTS

4.1. Evaluation of GCMs and downscaling methods

Although the GCMs were selected based on the work of Kamruzzaman *et al.* (2021), A comparison of the accuracy has been performed to evaluate the performance of all the models, as well as the downscaling methods, to the area under study. The observed climate variables have been used to calculate the NS, R^2, and RMSE for all the models and methods. Precipitation data has been obtained by downscaling the 'pr' variable to the IMD 2D grids. It is observed that MIROC6 has the least RMSE (107.19 cm) and highest R^2 (0.52) as compared to other GCMs. The prediction capability of INM-CM5-0 is also comparable to MIROC6 for an Indian semi-arid environment. Among all the methods used to downscale the precipitation, LS has proven to be more accurate for MIROC6. Although for other models, the accuracy is highly variable. The PT method has shown consistently high accuracy matrices for all models. In this study, the PT method has been utilised to downscale the precipitation values.

The temperature values were obtained by downscaling the 'tasmax' and 'tasmin' parameters of GCM, which correspond to the maximum daily temperature and minimum daily temperature, respectively. The prediction accuracy of the GCMs, along with the statistical methods, are consistently high in all the combination (GCM and downscaling method) (Tables 2–4). ACCESS-CM2 has shown slightly better accuracy with VS ($R^2 = 0.89$) method for maximum temperature. For minimum

Table 2 | Accuracy scores (average basin values) for different methods for the bias correction of downscaled **monthly precipitation** from different CMIP6 GCMs

	DM			LI			LS			PT		
	NS	R2	RMSE *(mm/month)*	NS	R2	RMSE *(mm/month)*	NS	R2	RMSE *(mm/month)*	NS	R2	RMSE *(mm/month)*
ACCESS-CM2	0.09	0.34	137.54	− 0.97	0.23	199.46	− 0.96	0.23	199.01	0.24	0.45	124.25
ACCESS-ESM1-5	− 0.07	0.33	147.01	− 1.04	0.24	203.08	− 1.04	0.24	203.21	0.00	0.37	141.68
INM-CM5-0	0.01	0.40	139.88	0.11	0.41	133.61	0.15	0.42	130.66	0.26	0.46	121.85
MIROC6	0.37	0.49	113.00	0.42	0.51	108.11	0.43	0.52	107.19	0.41	0.50	109.53

Table 3 | Accuracy scores (average basin values) for different methods for the bias correction of downscaled maximum monthly mean temperature from different CMIP6 GCMs

	DM			LS			VS		
	NS	R2	RMSE *(°C/month)*	NS	R2	RMSE *(°C/month)*	NS	R2	RMSE *(°C/month)*
ACCESS-CM2	0.87	0.88	1.81	0.82	0.83	2.17	0.89	0.89	1.70
ACCESS-ESM1-5	0.87	0.87	1.85	0.75	0.78	2.55	0.89	0.89	1.71
INM-CM5-0	0.87	0.87	1.84	0.81	0.83	2.21	0.88	0.88	1.75
MIROC6	0.85	0.86	1.96	0.56	0.66	3.36	0.86	0.86	1.90

Table 4 | Accuracy scores (average basin values) for different methods for the bias correction of downscaled minimum monthly mean temperature from different CMIP6 GCMs

	DM			LS			VS		
	NS	R2	RMSE *(°C/month)*	NS	R2	RMSE *(°C/month)*	NS	R2	RMSE *(°C/month)*
ACCESS-CM2	0.95	0.95	1.34	0.93	0.93	1.65	0.95	0.95	1.39
ACCESS-ESM1-5	0.95	0.95	1.34	0.93	0.94	1.59	0.94	0.94	1.50
INM-CM5-0	0.95	0.95	1.34	0.91	0.92	1.84	0.94	0.94	1.54
MIROC6	0.95	0.95	1.39	0.93	0.93	1.62	0.94	0.94	1.49

temperature, the accuracy achieved by DM with ACCESS-CM2 ($R^2 = 0.95$) outputs is slightly higher than other combinations. The accuracy for temperature prediction is relatively high for all the combinations, which is manifested by the seasonality of the data. All the models with the bias correction method are recommended for the downscaling of temperature except the LS method since the accuracy achieved is relatively low with respect to other methods (Table 3).

4.2. Impact on basin hydrology

The climate variables have been analysed for the deviation from the baseline scenario. The baseline values of the variables have been determined by predicting the climate data with the LSTM algorithm (Long-Short Term Memory), with observed daily precipitation and temperature data from 1975 to 2020. The deviation from the baseline climate has been analysed for all three scenarios (SSP1, SSP2 and SSP5). The SSP1 is characterised by more precipitation from 2030 to 2043; after that, there is a deficiency in rainfall. SSP2 and SSP5 have shown continuous deficiency of as much as 529 mm in rainfall from the baseline. The results suggest a decrease in the rainfall period, hence increasing the intensity of rainfall. The maximum temperature has shown a variable increase of 1.23 °C–3.16 °C for SSP1. The net change in temperature change for SSP2 and SSP5 is more or less equal, which has shown an increase in maximum temperature of 2.51 °C. The minimum temperature has shown a change of 2.94 °C, 2.51 °C and 2.50 °C for SSP1, SSP2 and SSP5, respectively. The minimum temperature has a steady increase (1.41–2.74 degrees from baseline) from 2030 to 2038 for SSP1. However, the rainfall has shown an increase in this time interval. The temperature rises from 2030 to 2038 manifest drought conditions in summer with high-intensity rainfalls that can bring flood in the rainy season.

Potential Evapotranspiration (PET) and Actual Evapotranspiration (ET) have been affected by the increased temperature in future scenarios. There is a decrease of 7.5%,4.3% and 5.75% in PET in the month of May and June for SSP1, SSP2 and SSP5, respectively, from baseline. The PET increases in the month of July, August and September by an average of 25%. The ET has shown considerable increase throughout the year, with an average increase of 53.6%, 51.2% and 48.5% for SSP1, SSP2 and SSP5, respectively, with the maximum increase in June, August and September. The increased transpiration process affects groundwater storage, which affects the lean period base flow to the streams. There is a decrease of 33% in lean period flow (October to May) in the Gopad river. The SSP2 manifests the most decrease of 42.8% in the river flow, followed by SSP5 (37.8%). The SSP1 has shown the slightest decrease in the river flow (21.6%) compared to the baseline scenario. The SWAT simulations indicate that there is a considerable decrease in groundwater recharge as compared to the baseline scenario. The results indicate an overall decrease of 12%, 25.3% and 23.3% in annual recharge to the groundwater for SSP1, SSP2 and SSP3, respectively. The volume of runoff generation has increased by more than 200% due to increased intensity of rainfall.

4.3. HCS

The average annual rainfall from 2021 to 2060 is found to vary from 540 mm to 1,628 mm spatially. The subbasin in the lower part of the Gopad river basin is projected to receive precipitation from 1,465 mm to 1,628 mm (annual average from 2021 to 2060). In response to projected precipitation, subbasins 19,22 and 26 have shown the highest surface runoff, followed by subbasins 9,27,34 and 53. The runoff generated by the subbasins in the upper Gopad basin has shown uniform values manifesting in the forest area. The groundwater recharge (shallow and deep) in subbasins 39–53 was the highest, creating a zone of high recharge potential. These subbasins are from the upper Gopad basin and are covered by forest and shrublands. These account for 56.55% of total groundwater recharge (Figure 7). The effect of high groundwater recharge comes with a high groundwater leakage to streams in these subbasins, which makes them critical areas for baseflow as they contribute 69.2% of total baseflow to the river. These areas also manifest high evapotranspiration (47% of the entire basin value) and high stream water yield (53.57% of the entire basin value). The subbasins 11,16,20,39,41,42,43,44,45,48,49,52 and 53 have been identified as HCS.

5. DISCUSSION

5.1. GCM and downscaling methods

The performance of four GCMs of CMIP6 in simulating rainfall and temperature for hydrological study has been evaluated from 1980 to 2020. The downscaling methods for these models have also been evaluated for their bias correction accuracy. MIROC6 with linear scaling has performed best for precipitation data, followed by INM-CM5-0 with the power transformation bias correction method. Accuracy is much more consistent for power transformation for all GCMs. Also, the variation of the accuracy for all the grid points, downscaled by power transformation, is much lower as compared to

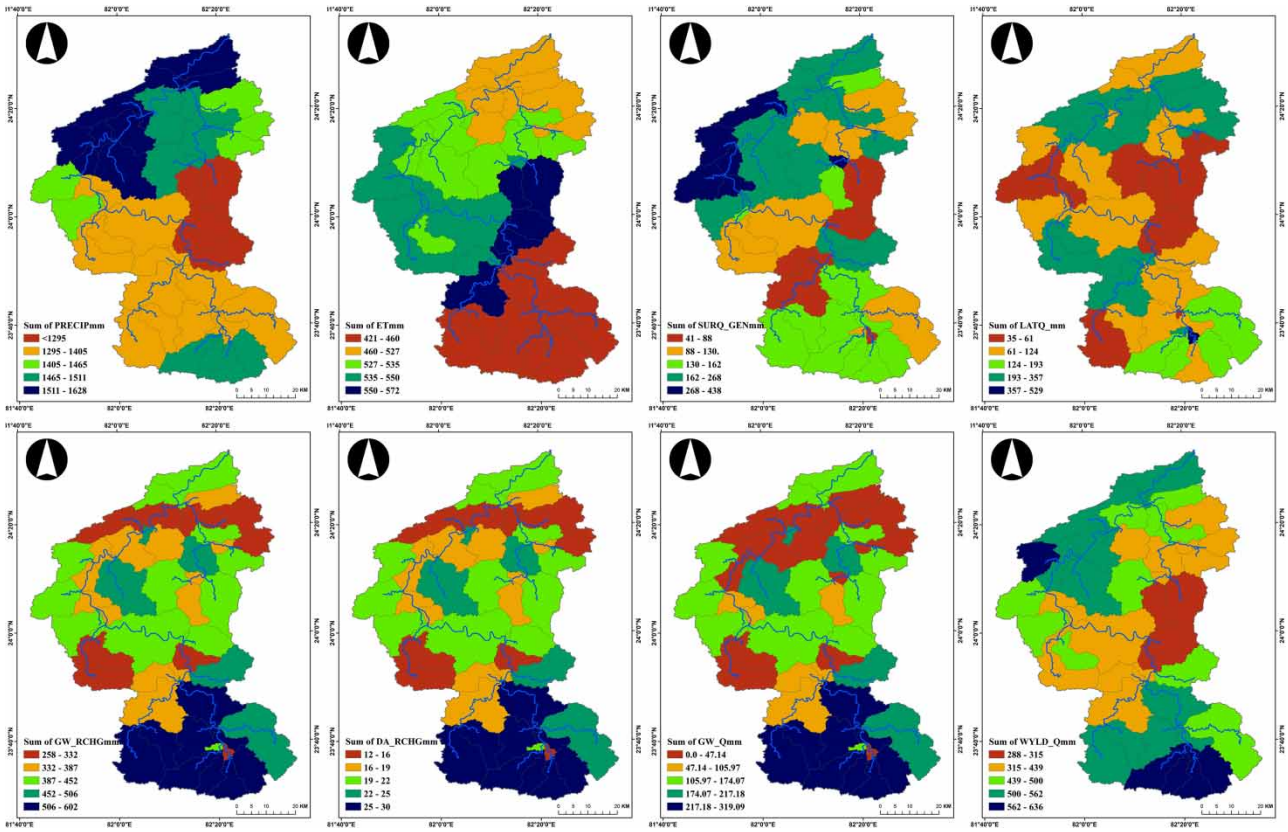


Figure 7 | Basin wise distribution of hydrological parameters (PRECIPmm-Precipitation in mm, ETmm- Evapotranspiration in mm, SURG_GENmm- Surface runoff generated in mm, LATQ_mm- Lateral flow to the streams in mm, GW_RCHGmm- Grounwater recharge in mm, DA_RCHGmm- Deep aquifer recharge in mm, GW_Qmm- Groundwater contribution to stream flow in mm, WYLD_Qmm- Water yield from subbasin to the streams in mm).

other methods. In this study, the power transformation has been used for the downscaling of MIROC6 precipitation data. As for temperature data, both variance scaling and distribution mapping have performed better than linear scaling. Variance scaling and distribution mapping have been used for downscaling maximum and minimum temperature, respectively. The overall accuracy in the simulation and downscaling of precipitation data has been the lowest as compared to temperature due to the lack of seasonality in the data series. The dry season precipitation simulated by the models was found to be higher. The ACCESS-CM2 was unable to capture the peak rainfall events as the simulated precipitations were found to be lower than the observed values. The MIROC6 and INM-CM5-0 performed better in simulating Indian weather as compared to others. As for the IPCC scenarios, SSP1 for precipitation has shown the highest increase in rainfall as compared to SSP2 and SSP5. For temperature, SSP1 and SSP2 have shown similar variations as compared to SSP5, which indicates a drastic increase in the temperature.

5.2. Future climate and hydrology of the Gopad river basin

The rainfall intensity has shown an increase in the near future. This could lead to higher runoff generation for the watershed in a short duration, causing flooding events in the lower Gopad basin (Figure 8). On the other hand, increased temperature in all the scenarios is a sign of high evapotranspiration, which can lower the groundwater table in the near future, causing a reduction in baseflow to the river. The increased temperature can reduce the flora of the upper Gopad basin, which further aggravates the problem of high runoff and low baseflow.

The SWAT model is calibrated and simulated with the bias-corrected future climate data. Surface water hydrology is found to be sensitive to climate variables. The peak discharge at the basin outlet has increased for all the IPCC scenarios. The dry

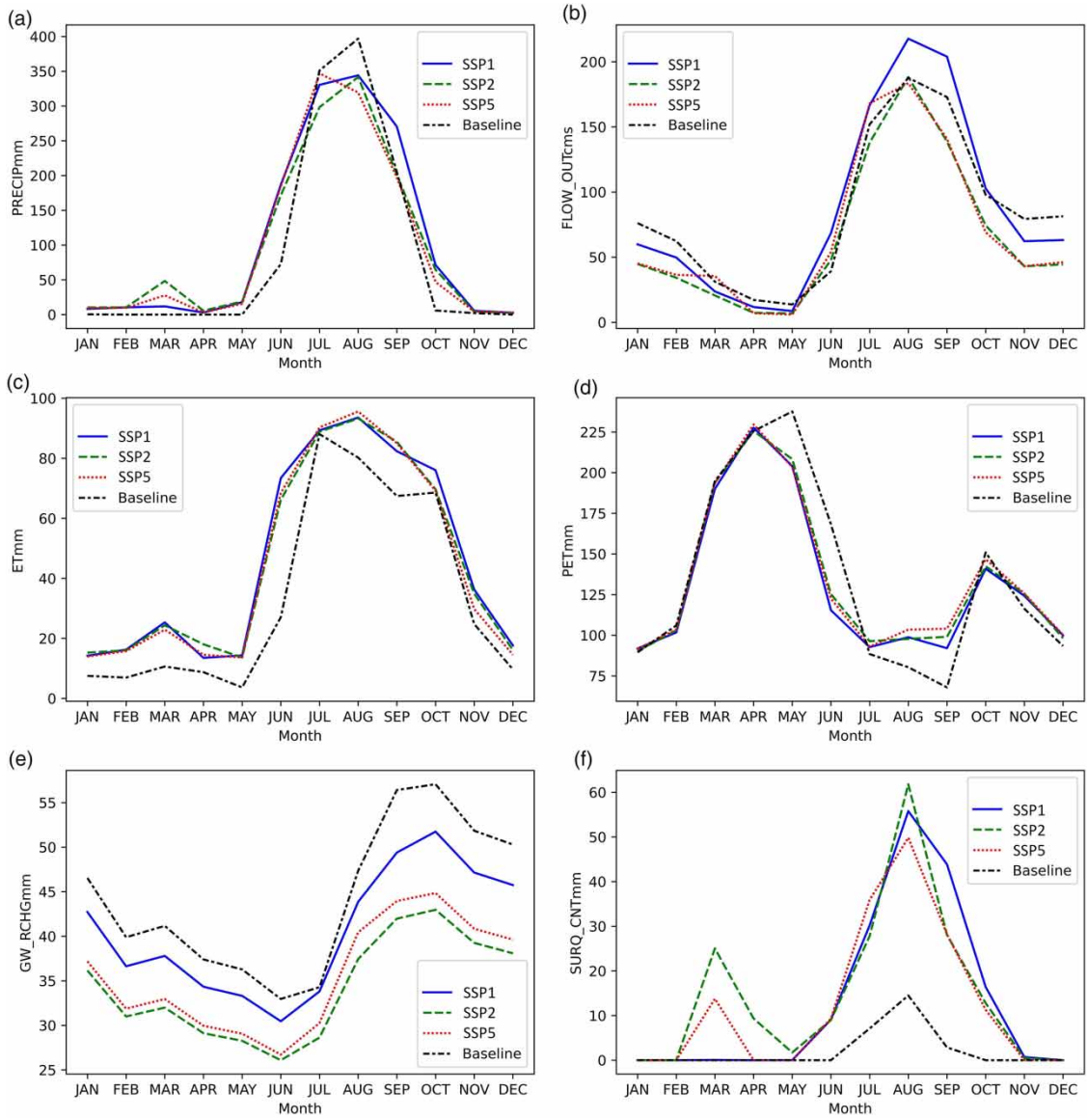


Figure 8 | Monthly average basin values of- (a) PRECIPmm (Precipitation), (b) FLOW_OUTcms (Stream discharge in m^3/sec) for the whole Gopad river basin. (c) ETmm (Evapotranspiration), (d) PETmm (Potential evapotranspiration), (e) GW_RCHGmm (Groundwater recharge) and (f) SURQ_CNTmm (Surface runoff contributing to stream flow). The values are represented for all three SSPs with the baseline.

season flow has shown a decline in discharge, accompanied by lower groundwater recharge. This is caused by the increased rainfall intensity in the basin, leading to a high runoff-to-infiltration ratio.

5.3. Hydrologically critical areas for future

The HCS has been delineated based on three ratios viz. runoff per unit rainfall, groundwater recharge per unit rainfall and groundwater leakage per unit stream flow. The critical subbasins for individual parameters have been identified by a threshold value determined by the 50th percentile of all basin values. The subbasins satisfying all three critical parameters have been identified as HCS (Figure 9). The spatial distribution of the HCS is high in the upper Gopad river basin (Figure 10). This area is highly pristine, with no human intervention, and covered with forest. The loamy soil in the area is highly permeable, leading to high groundwater recharge. The slopy topography of the sub-watersheds supports the groundwater leakage to the river; hence, these subbasins contribute significantly to the baseflow. The area of HCS is less than 40% of the total basin area, and these HCS contribute significantly to the total basin water budget. It is required to maintain the hydrogeological settings

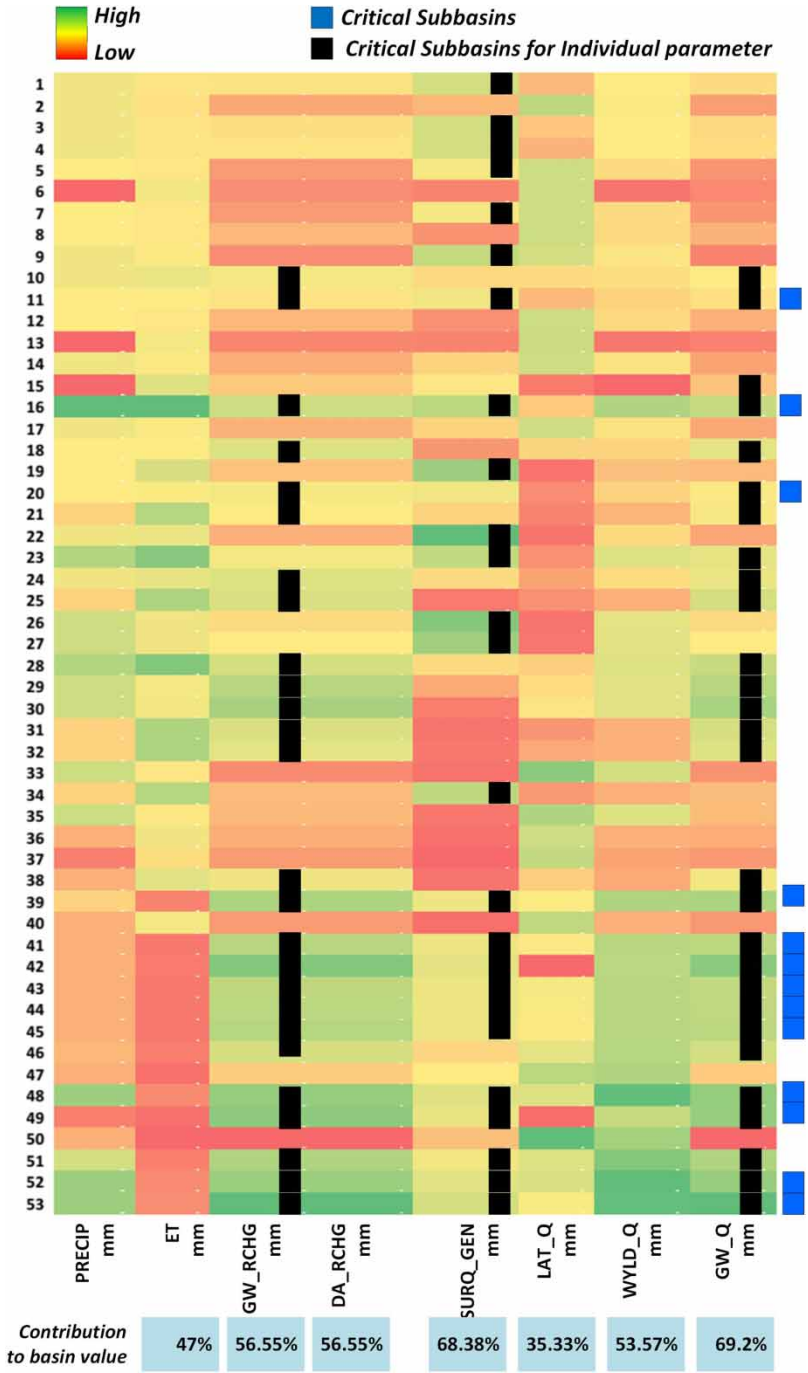


Figure 9 | Raster plot of the subbasin-wise variation of SWAT outputs. The black squares represent the subbasins with the variable's magnitude higher than the threshold for the given parameter (GW_RCHG, SURQ_GEN and GW_Q). The blue squares represent the overall critical subbasins. On the bottom, the contribution of the identified critical subbasins to the overall basin values has been tabulated. For example, the identified critical subbasins contribute 56.55% to the groundwater recharge compared to all subbasins in the study area.

in the area for the river to be perennial. The activities such as mining and reservoir construction can significantly change the river discharge.

The critical subbasins in the study area are mapped based on the ratio of runoff and groundwater recharge to total rainfall. The groundwater contribution to the stream flow has also been considered as a parameter. All the parameters have been derived from the SWAT model, which is highly sensitive to weather parameters. Therefore, the critical subbasins are sensitive

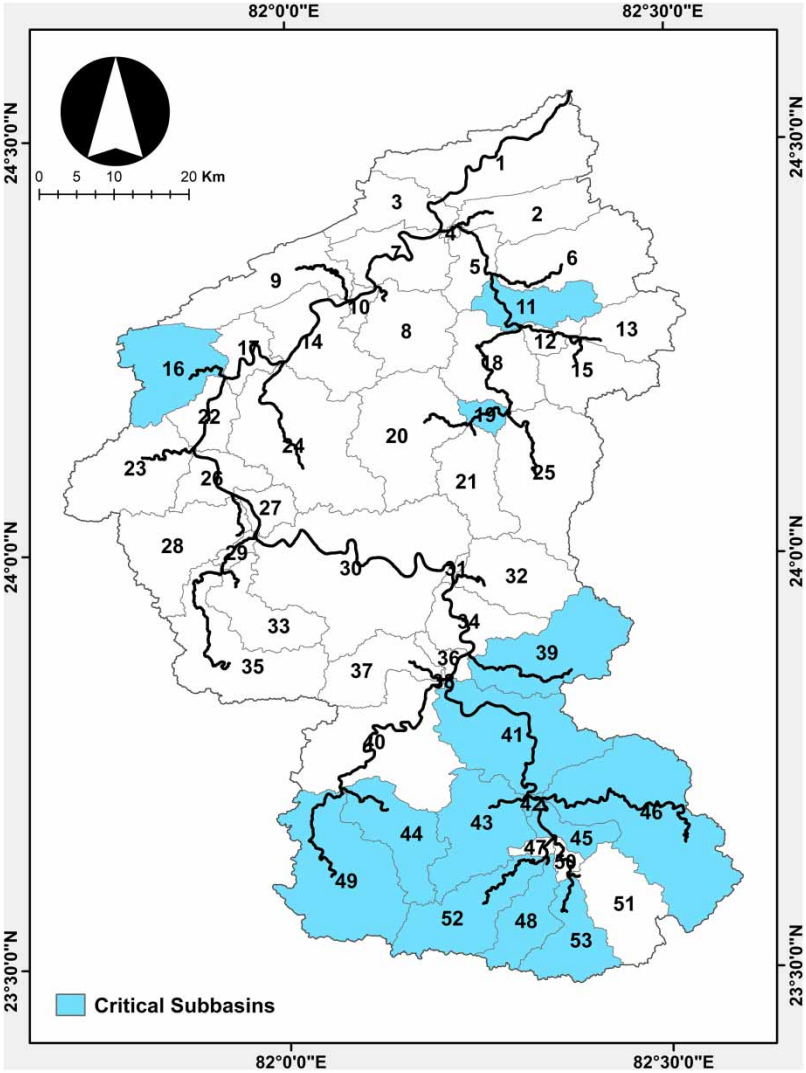


Figure 10 | Hydrologically critical subbasins in the Gopad river watershed.

to rainfall distribution, temperature and other weather parameters. The sensitivity of critical subbasins to the change in climate patterns should be analysed in future research work.

With water scarcity becoming increasingly prevalent in India, it is crucial to focus on conserving the existing hydrological balance of a region both presently and in the future. By using climate models, we can gain insights into future scenarios and analyse the potential impact of climate change on these delicate natural processes. The areas identified as critical zones in this research must be protected for water resource conservation purposes. In future research, it is important to identify critical zones for groundwater storage to ensure the holistic conservation of both surface and groundwater resources.

6. CONCLUSION

The hydrology of the Gopad river basin has been simulated by SWAT Model with the integration of GCMs outputs to identify the hydrologically critical subbasins for future IPCC scenarios. The climate change impact has been discussed along with a comparative analysis of the CMIP6 Models and the downscaling techniques. It is found that the MIROC6 and ACCESS-CM2 better simulate the precipitation and temperature for the Gopad river basin, respectively. In the statistical downscaling methods, the performances of DM and VS were superior w.r.t other methods. The precipitation in the monsoon season for 2021–2060 is projected to be 9.2% less than the baseline. At the end of 2060, the average annual temperature is projected

to rise by 1.23 °C–3.16 °C. Due to these changes in the climate, the peak flow in the Gopad river is projected to be increasing, and the lean flow is decreasing. The groundwater recharge is projected to fall by 12–25.3%, leading to a decreased baseflow. The critical subbasins for future scenarios have been identified to be primarily situated in the upper Gopad river basin. Most of the HCS has been identified in forest areas. It is suggested to prevent significant alteration to the hydrogeology and land use of these HCS for the better hydrological balance of the study area in the near future. It can be concluded that the hydrological factor affecting the critical subbasins are land use, soil infiltration capacity, topography and rainfall distribution. The sensitivity analysis of these parameters should be performed in future studies.

ACKNOWLEDGEMENTS

The authors would like to express their sincere thanks to Hindalco Industries Limited (Mahan, M.P., India) for funding the research and providing the river discharge data.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abbaspour, K. C. 2019 *SWATCUP-2019 SWAT Calibration and Uncertainty Programs*.

Arnold, J., Kiniry, R., Williams, E., Haney, S. & Neitsch, S. 2012 *Soil & Water Assessment Tool, Texas Water Resources Institute-TR-439*.

Chen, J., Brissette, F. P., Poulin, A. & Leconte, R. 2011 Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research* **47** (12). https://doi.org/https://doi.org/10.1029/2011WR010602.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. & Taylor, K. E. 2016 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. *Geoscientific Model Development* **9** (5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016.

Gusain, A., Ghosh, S. & Karmakar, S. 2020 Added value of CMIP6 over CMIP5 models in simulating Indian summer monsoon rainfall. *Atmospheric Research* **232** (June 2019), 104680. https://doi.org/10.1016/j.atmosres.2019.104680.

Kamruzzaman, M., Shahid, S., Islam, A. T., Hwang, S., Cho, J., Zaman, M. A. U., Ahmed, M., Rahman, M. M. & Hossain, M. B. 2021 Comparison of CMIP6 and CMIP5 model performance in simulating historical precipitation and temperature in Bangladesh: a preliminary study. *Theoretical and Applied Climatology* **145** (3–4), 1385–1406. https://doi.org/10.1007/s00704-021-03691-0.

Kumar, A., Singh, A. & Gaurav, K. 2022 Assessing the synergic effect of land use and climate change on the upper Betwa River catchment in Central India under present, past, and future climate scenarios. *Environment, Development and Sustainability*. (0123456789). https://doi.org/10.1007/s10668-022-02260-3.

Leander, R., Buishand, T. A., van den Hurk, B. J. J. M. & de Wit, M. J. M. 2008 Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output. *Journal of Hydrology* **351** (3), 331–343. https://doi.org/https://doi.org/10.1016/j.jhydrol.2007.12.020.

Maraun, D., Widmann, M. & Gutiérrez, J. M. 2019 Statistical downscaling skill under present climate conditions: a synthesis of the VALUE perfect predictor experiment. *International Journal of Climatology* **39** (9), 3692–3703. https://doi.org/https://doi.org/10.1002/joc.5877.

Mehta, V. K., Walter, M. T., Brooks, E. S., Steenhuis, T. S., Walter, M. F., Johnson, M., Boll, J. & Thongs, D. 2004 Application of SMR to Modeling Watersheds in the Catskill Mountains. *Environmental Modeling & Assessment* **9** (2), 77–89. https://doi.org/10.1023/B:ENMO.0000032096.13649.92.

Melillo, J. M., Richmond, T. T. & Yohe, G. 2014 Climate change impacts in the United States. *Third National Climate Assessment*. U.S. Global Change Research Program, Washington, DC. https://www.nrc.gov/docs/ML1412/ML14129A233.pdf .

Mohammed, I. N., Bomblies, A. & Wemple, B. C. 2015 The use of CMIP5 data to simulate climate change impacts on flow regime within the Lake Champlain Basin. *Journal of Hydrology: Regional Studies* **3**, 160–186.

Muto, Y., Noda, K., Maruya, Y., Chibana, T. & Watanabe, S. 2022 Impact of climate and land-use changes on the water and sediment dynamics of the Tokoro River Basin, Japan. *Environmental Advances* **7**, 100153. https://doi.org/10.1016/j.envadv.2021.100153.

Needelman, B. A., Gburek, W. J., Petersen, G. W., Sharpley, A. N. & Kleinman, P. J. A. 2004 Surface runoff along two agricultural hillslopes with contrasting soils. *Soil Science Society of America Journal* **68** (3), 914–923. https://doi.org/https://doi.org/10.2136/sssaj2004.9140.

Niraula, R., Kalin, L., Srivastava, P. & Anderson, C. J. 2013 Identifying critical source areas of nonpoint source pollution with SWAT and GWLF. *Ecological Modelling* **268**, 123–133. https://doi.org/10.1016/j.ecolmodel.2013.08.007.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K. & Sanderson, B. M. 2016 The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9** (9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S. & Mukhopadhyay, B. 2014 Development of a new high spatial resolution (0.25° X 0.25°)Long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. In: MAUSAM, 65th ed. pp.1–18.

Rivera, J. A. & Arnould, G. 2020 Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: climatic features and long-term trends (1901–2014). *Atmospheric Research* **241**, 104953. https://doi.org/https://doi.org/10.1016/j.atmosres.2020.104953.

Sharannya, T. M., Mudbhatkal, A. & Mahesha, A. 2018 Assessing climate change impacts on river hydrology – A case study in the Western Ghats of India. *Journal of Earth System Science* **127** (6), 1–11. https://doi.org/10.1007/s12040-018-0979-3.

Sun, X., Bernard-Jannin, L., Garneau, C., Volk, M., Arnold, J. G., Srinivasan, R., Sauvage, S. & Sánchez-Pérez, J.-M. 2016 Improved simulation of river water and groundwater exchange in an alluvial plain using the SWAT model. *Hydrological Processes* **30** (2), 187–202.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. & Hsu, K.-L. 2018 A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics* **56** (1), 79–107. https://doi.org/https://doi.org/10.1002/2017RG000574.

Tegegne, G., Park, D. K. & Kim, Y.-O. 2017 Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin. *Journal of Hydrology: Regional Studies* **14**, 49–66. https://doi.org/10.1016/j.ejrh.2017.10.002.

First received 10 December 2022; accepted in revised form 10 April 2023. Available online 21 April 2023