



Climate change impact on extreme precipitation and peak flood magnitude and frequency: observations from CMIP6 and hydrological models

Hadush Meresa^{1,2} · Bernhard Tischbein¹ · Tewodros Mekonnen²

Received: 16 March 2021 / Accepted: 28 November 2021 / Published online: 24 January 2022
© The Author(s) 2022

Abstract

Changes in climate intensity and frequency, including extreme events, heavy and intense rainfall, have the greatest impact on water resource management and flood risk management. Significant changes in air temperature, precipitation, and humidity are expected in future due to climate change. The influence of climate change on flood hazards is subject to considerable uncertainty that comes from the climate model discrepancies, climate bias correction methods, flood frequency distribution, and hydrological model parameters. These factors play a crucial role in flood risk planning and extreme event management. With the advent of the Coupled Model Inter-comparison Project Phase 6, flood managers and water resource planners are interested to know how changes in catchment flood risk are expected to alter relative to previous assessments. We examine catchment-based projected changes in flood quantiles and extreme high flow events for Awash catchments. Conceptual hydrological models (HBV, SMART, NAM and HYMOD), three downscaling techniques (EQM, DQM, and SQF), and an ensemble of hydrological parameter sets were used to examine changes in peak flood magnitude and frequency under climate change in the mid and end of the century. The result shows that projected annual extreme precipitation and flood quantiles could increase substantially in the next several decades in the selected catchments. The associated uncertainty in future flood hazards was quantified using aggregated variance decomposition and confirms that climate change is the dominant factor in Akaki (C2) and Awash Hombole (C5) catchments, whereas in Awash Bello (C4) and Kela (C3) catchments bias correction types is dominate, and Awash Kuntura (C1) both climate models and bias correction methods are essential factors. For the peak flow quantiles, climate models and hydrologic models are two main sources of uncertainty (31% and 18%, respectively). In contrast, the role of hydrological parameters to the aggregated uncertainty of changes in peak flow hazard variable is relatively small (5%), whereas the flood frequency contribution is much higher than the hydrologic model parameters. These results provide useful knowledge for policy-relevant flood indices, water resources and flood risk control and for studies related to uncertainty associated with peak flood magnitude and frequency.

Keywords CMIP6 · Bias correction · Hydrology · Extremes · Model structure · Uncertainty · Impact

1 Introduction

In the coming decades, climate change will likely become a primary problem affecting hydrological regimes and flood hazard conditions. According to the IPCC report, significant changes in atmospheric temperature, precipitation, humidity, and circulation are expected, resulting in increased extreme events, including floods, droughts, heatwaves, heavy rainfall, and more intense cyclones (Knutti and Sedláček 2013; IPPC 2013). According to Serdeczny et al. (2016) and Rojas et al. (2013), flood frequency will be increased due to climate change and will have a massive socio-economic impacts. Serdeczny et al. (2016) stated that there have been climate-linked changes in the magnitude, frequency and timing of flood characteristics in different regions of Africa. The autumn and winter flood are significantly increasing due to a large positive change in rainfall intensity and frequent during the autumn and winter seasons. In contrast, the decline in peak floods are due to decreasing future precipitation and increasing evapotranspiration (Coulibaly et al. 2020; Thober et al. 2018; Balke and Nilsson 2019; Blöschl et al. 2019). The understanding of catchment-scale flood hazard projections requires a chain of linked concepts and comprehensive framework. Therefore, developing a comprehensive cascade modeling that considers these principal components of climate change and hydrologic modeling framework is crucial.

Assessing the impact of climate change on local hydrological extremes is difficult, mainly due to the uncertainty in propagating information from a global coarse resolution model to the local scale. Meresa and Romanowicz (2017) studied the critical uncertainty embedded (hydrological parameter uncertainty (HBV), climate models (RCP), and distribution parameter uncertainty (Generalize Extreme Value (GEV)) in the projection of hydrological extremes and addressed that local-scale impact study is highly prone to climate models chain error. Joseph et al. (2018) also considered hydrological parameters (VIC) and climate models (RCP) uncertainty assessment in seasonal flow projections. Soriano et al. (2019) evaluated the uncertainty of bias correction methods on flood frequency and found that the climate model's bias correction is significant and brought significant change in the magnitude of flood design values. Feng and Beighley (2020) also identified the uncertainties from both climate forcings and hydrologic model components as well as associated parameterization. They compared the contribution of climate models, hydrological model structure and hydrological parameters using the seasonal flow indices as hydrological indicators. For the changes in seasonal mean flow, hydrologic models and climate models are two main uncertainty contributors. In contrast, the role of hydrologic parameters is not significant in seasonal flow and high flow (Meresa et al. 2021). The uncertainty in future climate flood hazard assessment arises from the uncertainties in the different component elements of the modeling chain used to calculate streamflow from model-derived estimates of future rainfall. These include uncertainties in the model-estimated parameters of future rainfall and temperature, climate downscaling approaches for calibrating meteorological conditions at a particular basin, the structure and parameters of the hydrological model of streamflow, and the flood frequency model used to assess flood recurrence periods (Bastola et al. 2011; Charles et al. 2019; Meresa 2020; Lawrence 2020; Meresa et al. 2021). Therefore, it is crucial to estimate and understand the complex and uncertain impact and challenging to identify the primary sources of uncertainty in future hazard estimation.

In the other hand, different climate projects (SRES, CORDEX, CMIP) have employed different ensembles of regional climate model (RCMs) and global climate models (RCMs) with different atmosphere-land modules to derive a wide range of possible future climate

conditions. Therefore, the locally projected hydrological extremes are highly dependent on the spread (or range) of forcing conditions from the ensemble of the RCM/GCM estimates employed in the local impact study area (Woldemeskel et al. 2014; Meresa and Romanowicz 2017; Hattermann et al. 2018; Charles et al. 2019; Meresa 2020; Lawrence 2020). The information from the spreads of multiple GCMs is considered as a band of impact uncertainty and quantified using the variance of the climate models (Meresa and Romanowicz 2017; Lawrence 2020). Bias correction methods of GCMs output are also essential and significantly impact the projected hydrological extremes magnitude and direction (Pierce et al. 2015; Hattermann et al. 2018; Charles et al. 2019). Similarly, the biases correction of GCMs output was applied on a grid commensurate with daily climate model output. Nowadays, many researchers have been used and addressed the importance of climate bias correction methods (Kay et al. 2009; Saini et al. 2015; Soriano et al. 2019). Quantile mapping, empirical mapping, simple statistical transformation, and joint probability are the most popular and widely applied for climate model bias correction and related to observed climate time series. These techniques will reduce the original climate model variability and outliers and prepare for forcing hydrological models. From the above study, the major source of uncertainty in climate projection is explained by selecting bias correction methods and the spread of selected GCMs. Furthermore, uncertainty is also introduced through the structure and parameterization of hydrological models. Hydrological models translate the meteorological conditions of rainfall and temperature into predicted stream flows using parameterizations of varying complexity, which must be calibrated with measurement databases. Different hydrological models may show varying skill levels in simulating different kinds of flow regime, and this is important for assessing future climate-linked flood extremes (for e.g., Pechlivanidis et al. 2017; Joseph et al. 2018; Her et al. 2019). Of which, many of the studies attempt to address the uncertainty of the hydrological parameters in climate change impact study using GLUE (Beven and Binley 1992), multi-objective function optimization approach (Dakhlaoui et al. 2017; Zhang et al. 2019), Bayesian Model Averaging (Beigi et al. 2019). This implies that on top of the spread of climate models and bias correction methods, the hydrological parameter uncertainty using GLUE is also important in climate change impact on hydrological extremes. In addition to these sources of uncertainty, the flood frequency associated uncertainty is also important. Collet et al. (2017), Kay et al. (2009), Meresa and Romanowicz (2017), Lawrence (2020) found that the flood frequency distribution models are highly dependent on the distribution of extreme floods, the number of events, and the size and shape of the catchment. This shows that a significant effort has been made to evaluate climate change impacts on extreme hydrology, mostly focusing on identifying trends in extreme low flow and high flow magnitude with their associated uncertainty estimation. However, there is little attention to uncertainty and impact estimation in projected peak flood magnitude and frequency. This study applies comprehensive uncertainty and impact estimation of climate change on flood hazard approach.

This investigation has therefore been structured to evaluate the CMIP6 climate data set to assess hydrological extreme conditions at selected catchments in the Awash basin of Ethiopia. The location is a particularly effective test site with a large historical meteorological and hydrological data archive and is of particular interest because of large extremes. This includes to (i) Compare different bias correction techniques for climate change impact assessment, (ii) Estimate the hydrological structure and parameter uncertainty, (iii) Estimate the impact and uncertainty of climate change and bias correction approach on flood projections, and (iv) Identify the source of uncertainty associated with projected flood hazards.

Table 1 Statistical description of the selected gauged stations

Code	C1	C2	C3	C4	C5
River/lake	Awash_K	Akaki	Kela	Awash_B	Awash_H
Gauge at site at/Nr	U/S of Koka	Aba S	Welen	Bello	Melka H
Latitude [degree]	8.4	8.75	9.0	8.85	8.38
Longitude [degree]	39.02	38.72	38.27	38.42	38.78
Catchment Area [km ²]	7093.9	1475.8	67.5	2568.75	7656
Elevation [m]	1602	1994	2129	2060	2300
5% of river flow value [m ³ /s]	0.78	0.03	0.06	0.05	0.01
95% of river flow value [m ³ /s]	154.43	5.86	43.02	63.76	33.29
Mean flow[m ³ /s]	41.69	1.33	8.14	25.16	8.23
Mean precipitation [mm]	3.16	2.37	2.73	2.59	2.37
Annual max precipitation [mm]	132.6	120.9	89.2	67.3	77.9
Base flow index [-]	0.65	0.02	0.16	0.49	0.09

2 Study area and hydro-climate datasets

Five sub-catchments were selected for climate change impact assessment. These are located in different geographical and hydro-climatic regimes (Table 1 and Fig. 1) and inside the Awash basin study area. Awash_K (@U/S of Koka), Akaki (@ Aba S.), Kela (@ Welen), Awash_B (@ Bello), and Awash_H (@ Melka H) catchments are most representative for flood hazard estimation under varying climate conditions. The drainage area of the selected catchments ranges from 67.5 km² for Welen to 7093.9 km² for U/S of Koka. The elevation of these catchments varies from 1602 m (@ U/S of koka) to 2300 m (@ Melka H). The historical records of daily precipitation, temperature, and streamflow data were taken from the Ethiopian meteorological and hydrological institute for the period 1981–2010. The selected stations are located inside the Awash basin study area. The highest observed flow rates from the daily time series for the basins ranges from 82 m³/s (@ Welen) to 884 m³/s (@ U/S of Koka), which is proportional to the catchment area. Simultaneously, the annual maximum precipitation (or the 95% of annual precipitation) is not proportional to the catchment area and annual maximum flow (or 95% of annual flow). The high mean of surface runoff is directly related to the high coefficient of variance in river flow (e.g., U/S of Koka).

3 Modeling and numerical experiments

The impacts and uncertainty of projected flood hazard for selected Awash catchments were examined using a new climate dataset (CMIP6), three bias correction techniques (statistical quantile factor (SQF), distribution quantile mapping (DQM), and empirical quantile mapping (EQM)), four hydrological models (HBV, SMART, NAM and HYMOD) with GLUE (generalize likelihood uncertainty estimation)—Monte Carlos simulation of 30,000 hydrological parameter sets and general extreme value (GEV)) (Fig. 2). In this study, four stages of numerical modeling experiments have been performed. The investigation started by selecting natural and/or semi-natural catchments and then extracting, correcting, and evaluating new climate datasets (CMIP6) using the three statistical techniques. A sampling of

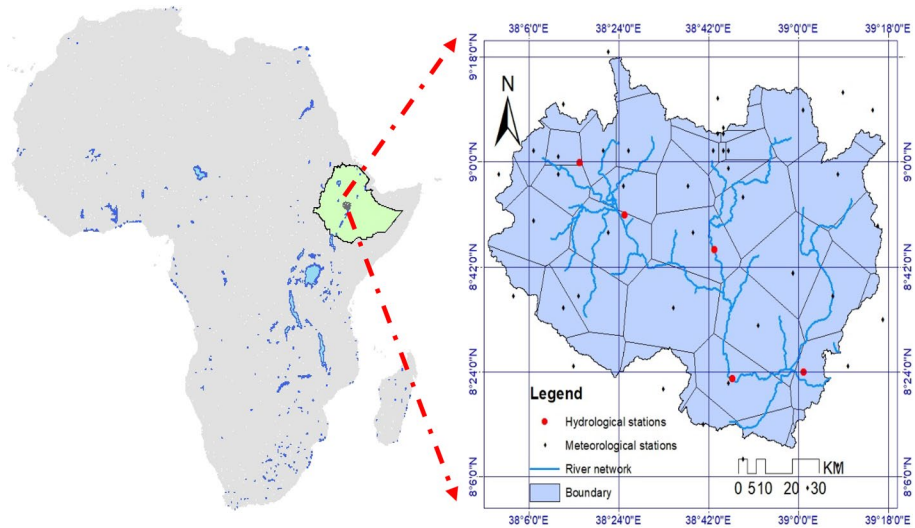


Fig. 1 Location of the selected case study. The red dots are the main hydrological stations chosen in this study

30,000 hydrological parameters set using Monte Carlo simulation (MCS) for hydrological parameter uncertainty band estimation using GLUE (Beven and Binley 1992), simulation of flood hazard projections using HBV, SMART, NAM and HYMOD conceptual hydrological models for hydrological model structure and parameter impact analysis. Finally, generalized extreme value (GEV), log normal (LNOR), Pearson-III (PEAR-II) was applied to estimate the extreme flood hazard magnitude and frequency (Fig. 2).

3.1 Coupled model intercomparison project phase 6 (CMIP6)

CMIP6 is a new climate dataset project (<https://pcmdi.llnl.gov/CMIP6/>) which was released by a collaborative effort of different climate research institutes to advance climate change knowledge and applicability. Precipitation and air temperature daily time series were extracted from the CMIP6 simulation datasets (<https://esgf-node.llnl.gov/search/cmip6/>). The climate models output covers 1971–2100 time period and provides global scale spatial resolution, ranging from 50 to 250 km. The three shared socioeconomic pathways (SSP) scenarios used in CMIP6 are SSP126 (lower level), SSP370 (middle level), and SSP585 (high level). In this study, 12 climate models and three CMIP6 scenarios (SSP126, SSP370, SSP585) were used to evaluate the projected flood hazard's impact and uncertainty in the selected Awash catchments (see Table S1 for climate model name and model details). The catchment centroid points were bracketed by four grid points of the coarse resolution GCMS, and the air temperature and precipitation from these points were extracted for further analysis. The climate models listed in Table S1 are all global models with 50–250 km resolutions. The CORDEX project has a fine resolution than CMIP6 and is assumed to simulate better than coarse resolution climate datasets. However, these GCMs can capture the variability and seasonality of precipitation and temperature better than those with high resolution (Almazroui 2019).

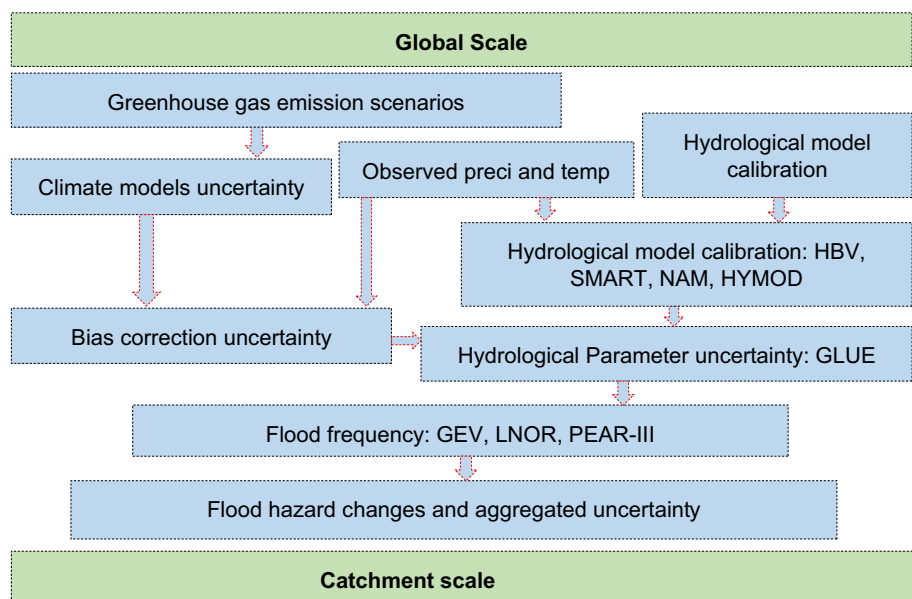


Fig. 2 Research flow chart to estimate projections of flood hazard and identify associated uncertainty

3.2 Climate projections—bias correction techniques

Global and regional climate models (GCM/RCMs) are essential for enlightening our understanding of annual, seasonal, and daily precipitation and air temperature characteristics. Climate model predictions differ from each other because of differences in the way fluxes are parameterized and as well as the underlying grid structure. This gives rise to biases in their outputs (Giorgi and Gao 2018; Krinner and Flanner 2018), making it challenging to develop accurate and reliable climate information. It has long been recognized that climate model bias must be corrected or ‘calibrated’ before use in other applications like the hydrological extremes projections (e.g., Ehret et al. 2012; Teng et al. 2012; Meresa and Romanowicz 2017; Osuch et al. 2017) in hydrology and water resources. However, minimizing the bias in climate model output is still a big challenge. There is a problem finding/developing a robust climate bias correction technique that reduces the biases between the observed and simulated climate data in the historical period. Previous studies have had mixed success in the application of different bias correction methods for precipitation (including intensity and wet days) and temperature data. For instance, Teutschbein and Seibert (2013), Yang et al. (2010) showed the distribution mapping based on theoretical distributions outperforms other bias correction methods in their result. Similarly, Berg et al. (2012), Chen et al. (2013) showed that theoretical distribution mapping performs similar to, or only marginally better than, empirical quantile mapping or other statistical methods. Contrary, Gudmundsson et al. (2012), Gutjahr and Heinemann (2013), Lafon et al. (2013) show that empirical quantile mapping demonstrates higher skill than theoretical distribution mapping in systematically correcting RCM precipitation. In this study, we used three climate bias techniques to correct the raw climate output and examine the contribution of bias correction methods to the flood hazard magnitude and frequency.

3.2.1 Statistical quantile factor (SQF)

This study proposed a new bias correction technique to compare with the commonly used methods (DQM and EQM). The statistical quantile factor change technique (SQM) of bias correction is a direct and straightforward precipitation correction, considered a simple error transfer from the historical to a future period. This involves correcting the future daily precipitation ($P_{\text{fut,corr}}$) by multiplying the sum of average ratio and reciprocal ratio of observed precipitation (P_{obs}) and reference precipitation simulation ($P_{\text{ref,raw}}$) to the prediction of raw climate model precipitation ($P_{\text{fur,raw}}$). By contrast, future air temperature is corrected ($T_{\text{fut,corr}}$) with an additive/subtractive constant applied to the raw climate output ($T_{\text{fur,raw}}$) and minimum difference of the raw ($T_{\text{ref,raw}}$) and observed (T_{obs}) air temperature in the reference period.

$$P_{\text{fut,corr}} = \frac{P_{\text{fur,raw}}}{2} * \left[\frac{P_{\text{obs}}}{P_{\text{ref,raw}}} + \frac{P_{\text{ref,raw}}}{P_{\text{obs}}} \right] \quad (1)$$

$$T_{\text{fut,corr}} = T_{\text{fur,raw}} + \min((T_{\text{obs}} - T_{\text{ref,raw}}), \pm 2.2) \quad (2)$$

3.2.2 Distribution quantile mapping (DQM)

Distribution quantile mapping is a parameter bias correction technique that depends on the type of distribution fitted to observed and simulated climate data (Piani et al. 2010). This distribution-based approach can be single or double distribution quantile mapping (S/DQM) or other distributions. The excess number of dry days, drizzles and wet days were considered and properly corrected in this method. For every N year, the zero precipitation is removed, and the single Gamma distribution is fitted to the upper 75% of daily precipitation. In contrast, the double Gamma distribution fitted to both upper and lower parts of 75% of the daily precipitation.

$$P_{\text{corr}} = F_{dg}^{-1}(F_{dg}(P_{\text{raw}}(t), \alpha_{\text{raw}}, \beta_{\text{raw}}), \alpha_{\text{Obs}}, \beta_{\text{Obs}}) \quad (3)$$

$$T_{\text{corr}} = F_{dn}^{-1}(F_{dn}(T_{\text{raw}}(t), \alpha_{\text{raw}}, \beta_{\text{raw}}), \alpha_{\text{Obs}}, \beta_{\text{Obs}}) \quad (4)$$

where P_{corr} and T_{corr} stand the bias-corrected daily precipitation and temperature, respectively. Likewise, P_{raw} and T_{raw} represent raw climate output, daily precipitation, and air temperature, respectively. The raw climate output inverse cumulative density (CDF) is symbolized by F_{dg} and F_{dn} for precipitation and temperature, respectively. The dn and dg subscripts stand for the normal (for temperature) and Gamma (for precipitation) distributions. The Gamma (for precipitation) distributions have two parameters, shape and scale parameters are symbolized by α and β , and the normal (for temperature) distribution standard deviation and mean are symbolized by σ and μ , respectively.

3.2.3 Empirical quantile mapping

Unlike the distribution mapping approach, the empirical quantile mapping is based on a paired-wise comparison between empirical cumulative density function (ecdf) of observed and simulated precipitation at the reference period. This is purely empirical and directly

matching the observed histogram to the future period. The future precipitation and temperature are corrected using the inverse of ecdf (ecdf^{-1}) and fitted ecdf

$$P_{\text{hst},m,d}^{\text{cor}} = \left(\text{ecdf}_{\text{obs},m}^{-1} \left(\text{ecdf}_{\text{hst},m} (P_{\text{hst},m,d}) \right) \right) \quad (5)$$

$$T_{\text{hst},m,d}^{\text{cor}} = \left(\text{ecdf}_{\text{obs},m}^{-1} \left(\text{ecdf}_{\text{hst},m} (T_{\text{hst},m,d}) \right) \right) \quad (6)$$

3.3 Bias correction performance evaluation

The performance of the selected statistical bias correction technique was evaluated using four statistical measures: correlation coefficient (RR), mean absolute error (MAE), root mean square error (RMSE), percent of bias (PBIAS), and maximum precipitation weighting root means square error (PWRMSE). These statistical measures were applied to daily rainfall and temperature time series of climate model output. The time series-based evaluation was performed by comparing the capacity of each approach to generate precipitation and temperature.

$$\text{RR} = \frac{\sum (P_s - P_s^-) * (P_c - P_c^-)}{\sqrt{\sum (P_s - P_s^-)^2 * \sum (P_c - P_c^-)^2}} \quad (7)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |P_s - P_c|}{N} \quad (8)$$

$$\text{PBIAS} = \frac{\sum_{i=1}^n (P_s - P_c)}{\sum_{i=1}^n P_s} \quad (9)$$

$$\text{PWRMSE} = \left[\frac{\sum_{i=1}^n (P_c - P_c)^2 * \left(\frac{P_s + \bar{P}_s}{2 * P_s} \right)}{N} \right]^{0.5} \quad (10)$$

where P_s is observed precipitation at a given station, P_c is corrected precipitation, N number of observations.

3.4 Hydrological modeling

Many hydrological models have been developed to understand hydrological processes on local and global spatial scales. There are vast ranges of hydrological model types; most of the hydrological models have considered physical mechanism and empirical equations to describe the catchment processes. The hydrological parameter set is one type of uncertainty source in extreme hydrological modeling. Precipitation-runoff models are often made up of a series of equations that describe a simplified description of hydrological processes at the small to large catchment scale. Lumped models treat the catchment process as a single unit (Keith 2001), with all variables being averaged over it at a given time scale.

This study looks at three different hydrological models with various structures. Each model gets forcings of catchment average precipitation and potential evapotranspiration input time series data at daily time-steps. At daily intervals, the models estimate actual evapotranspiration, change in stored water in various components, including the slow and quick runoff.

3.4.1 Hydrologiska Byråns Vattenbalansavdelning (HBV)

The HBV (Bergström 1976) is widely applied in different hydro-climate conditions of the world (Meresa and Gatachew 2019; Meresa et al. 2017; Her et al. 2019). The model is mainly designed to simulate streamflow using precipitation, temperature, and evapotranspiration (estimated using Hamon 1964) climate variables. It is a daily lumped conceptual hydrological model with fourteen parameters and applied in many parts of the world (Meresa et al. 2017; Meresa and Gatachew 2019). In this study, HBV was used with eleven hydrological parameters (Table S2). HBV model structure has three consecutive stores: one related to the surface, the second associated with the saturated soil layer, and the other to the unsaturated routing store. Daily precipitation and daily evapotranspiration (estimated using Hamon 1964) are the primary climate variable that are used as inputs to the model. The detailed information about the model is described in Bergström (1976). The upper and lower limits of the three hydrological models are listed in Table S2. The details of these hydrological models, including their physical meaning and description of each model parameter, are explained Bergström (1976).

3.4.2 Nedbr-Afstrmnings-model (NAM)

The NAM (Nielsen and Hansen 1973) is widely used all over the globe. It has been used to estimate the direct and runoff and investigate the contributions of groundwater and surface water to streamflow in catchments (O'Brien et al. 2013). The NAM model structure applied in this study has two main storage reservoirs for soil moisture accounting, including runoff routing and reservoirs, to characterize four hydrological process pathways. The model has eight parameters to govern the moisture content in both storages representing the surface runoff, soil and groundwater storages, and three parameters conveying to the runoff routing components (Table S2).

3.4.3 Soil moisture analysis rainfall tool (SMART) model

The SMART model (Mockler et al. 2016, 2014) was created as a hydrological model for water quality simulations. Individual contributions to hydrological flow paths are highlighted in the model. The model has ten parameters to control the precipitation-runoff process at a catchment scale (Table S2). Following the approach of SMARG and its predecessors, SMART simulates hydrological flows utilizing conceptual soil moisture accounting equations based on a number of soil moisture layers. The depths of these soil moisture layers vary based on the characteristics of the watershed, and they represent average conditions across the catchment. The four conceptual flow paths' numerical procedures involve a total of ten parameters, four of which are related to flow routing. The model, like the SMARG, comprises six equal-depth soil layers and six soil moisture accounting parameters (Table S2). Because drain flow can be an important conduit for nutrients in agricultural catchments (e.g., Madison et al. 2014), it is included as a separate flow path in the model. It is related to soil moisture surplus and the drain parameter (S), fluctuating between 0 and

1. The soil outflow coefficient is used to determine interflow, a mix of soil moisture excess and outflow from the soil layers (D). Individual outflow equations also related to the outflow coefficient (D) parameter, and used to compute shallow and deep groundwater at each subsurface layer.

3.4.4 HYdrologic MODel (HYMOD)

The HYMOD is a conceptual model that is lumped the hydrological cycle components together (Boyle 2000; Sun et al. 2010). The rainfall excess and two parallel sets of linear tanks are the two fundamental components of this scheme. It was decided to use a modified version of the model. The model now includes a snow module based on degree days. The degree-day factor DD, the precipitation/degree-day relation (Dew), and the threshold temperature are the three factors that control the melting storage. The maximum storage capacity (Cmax) and the degree of geographic variability of soil moisture capacity within the watershed determine soil moisture content. Excess water from the soil zone flows into quick-flow tanks and groundwater based on a partitioning factor that divides the flow between fast and slow reservoirs ($v1$ and $v2$). The water is dispersed into three linear reservoirs in series, one for the quick runoff component and one for groundwater flow in parallel with the other flows. The residence times R_s and R_q are used to classify reservoirs (Table S2).

3.5 Hydrological model parameter selection and evaluation

There are different ways of parameter sampling from the upper and lower boundary of hydrological parameters, which depend on the computing times and number of parameters in a specific hydrological model. Beven and Binley (2014) stated that there is no fixed threshold in parameter sampling that varies from ten thousand to a hundred thousand parameter sets. In this study, 30,000 parameter sets were generated from three hydrological model parameters range (Table S2). Characteristics for high flow and low flows have different and need two parallel calibration techniques (Meresa and Romanowicze 2017). Meresa and Romanowicze (2017) stated that the NSE likelihood function is relatively useful for high flow simulation due to the interest in peak flow and LogNSE for low flow. This study used the NSE, LogNSE, and KGE objective functions to simulate high flow regimes and evaluated against observed streamflow. Based on the model performance, 2000 sets of hydrological model simulations were selected as behavioral conditions.

$$NSE = 1 - \frac{\sum_{t=1}^j (Q_{o,t} - Q_{m,t})^2}{\sum_{t=1}^j (Q_{o,t} - \bar{Q}_o)^2} \quad (11a)$$

$$LNSE = 1 - \frac{\sum_{t=1}^j (\ln(Q_{o,t}) - \ln(Q_{m,t}))^2}{\sum_{t=1}^j (\ln(Q_{o,t}) - \ln(\bar{Q}_o))^2} \quad (11b)$$

where $Q_{o,t}$ and $Q_{m,t}$ is the observed and simulated flow at time t , Q_o is the mean observed flow, and j is the length of the j th time series.

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (11c)$$

$$\text{with } \beta = \frac{Q_{\text{obs}}}{Q_{\text{sim}}}; \alpha^2 = \frac{Q_{\text{sim}}^2}{\sigma_{\text{obs}}^2}; r = \frac{1}{n} \sum_{i=1}^n \frac{(Q_{\text{obs}} - \bar{Q}_{\text{obs}})(Q_{\text{sim}} - \bar{Q}_{\text{sim}})}{\sigma_{\text{obs}} \sigma_{\text{sim}}}.$$

Where Q_{obs} and Q_{sim} are the observed and simulated flows, σ_{obs} and σ_{sim} are, respectively, the standard deviation of the observed and simulated flows. The best values of each objective function were selected for hydrological model structure and parameter uncertainty analyses.

The hydrological model runs using the entire space of hydrological parameters combination and evaluated using goodness-of-fit criterion (Beven 2007). Assume likelihood function $H(X)$ to separate the non-behavioral and behavioral simulations produced by different variables X , such as input data, hydrological model parameters, hydrological model structures, and extreme frequency models. Every i th values of the variable X has its own one likelihood measure at time t . The ensemble of each variable X_i ($i = 1 \dots m$) provides the multi-likelihood measure values $H(X_i)$. The generalized likelihood uncertainty estimation (GLUE) function is shown in Eq. (12). The std/variance of residual σ^2_e value is the error in the estimated results affected by the model parameters/input data/hydrological model type/extreme frequency distribution models. If the estimated value of σ^2_e is near equal to the estimated maximum likelihood or equal to the std/variance of the observation data σ^2_o , the likelihood measure $H(X)$ is equal to zero, which indicates extremely high uncertainty.

$$H(X) = 1 - \frac{\sigma_o^2}{\sigma_e^2} \quad (12)$$

In this study, GLUE approach was used to estimate the uncertainty associated with hydrological parameters.

3.6 Flood frequency analysis

Extreme frequency analysis is crucial to understand the probability of reoccurrence of flood events at different return periods. This information plays an essential role in flood control and water resource planning. However, high flow frequency mainly depends on the frequency distribution model type and number of distribution parameters. Depending on the environmental and climatic background, many distributions are deployed in various countries by many researchers to estimate the frequencies of high flow. For example, the Log-Pearson III distribution model is popular in the USA and Australia for infrastructure design (Griffis and Stedinger 2007), GEV and PEAR-III in Europe (Refsgaard et al. 2013) and Africa (Meresa and Gatachew 2019), and Wakeby and Log-Normal distribution types have been frequently used in Asian countries (Chen et al. 2012). However, one or two statistical distribution models may not capture the entire temporal and spatial variability of hydrological extremes. Therefore, the most commonly used distribution type (GEV) was applied for flood hazard frequency and magnitude curve development. The distribution was fitted to peak flow to understand the hydrological flood in the selected Awash catchments. In Eq. (13), the probability density function (PDF) is presented and has three parameters.

$$\text{GEV } f(x) = \begin{cases} \frac{1}{\sigma} \exp\left(-\left(1 + kz\right)^{-\frac{1}{k}}\right) (1 + kz)^{-1-\frac{1}{k}} & k \neq 0 \\ \frac{1}{\sigma} \exp(-z - \exp(-z)) & k = 0 \end{cases} \quad (13)$$

where α is scale parameter, β shape parameter and k location parameter.

3.7 Uncertainty decomposition and estimation

In this study, three main variables, climate models, bias correction techniques, and hydrological parameters, were used to identify the relative uncertainty contribution on the total flood hazard magnitude. Unlike additive or multiplicative uncertainty estimation methods, ANOVA can decompose the aggregated source of uncertainty into individuals' and their interaction using specific extreme flow indices (Meresa and Romanowicz 2017). Three sources of uncertainty in flood hazard projections drive from climate models, bias correction methods, and hydrological parameters were estimated using the variances decomposition approach (ANOVA). In this study, n-way of ANOVA was used to distinguish the main variable effects and their interaction effect on the aggregated extreme frequency indices.

$$SST = SS_{CM} + SS_{BC} + SS_{HP} + SS_{CMBC} + SS_{CMHP} + SS_{BCHP} + \text{Error} \quad (14)$$

where SS_{CM} is the sum of the standard error of the climate models, SS_{BC} is the sum of the standard error of the bias correction methods, SS_{HP} sum of the standard of the hydrological parameters, SS_{CMBC} is sum standard error of combined effect of climate models and bias correction methods, SS_{BCHP} sum standard error of combined impact of climate bias correction methods and hydrological parameters.

4 Results

4.1 Evaluation and validation of different climate bias correction techniques

The annual and seasonal maximum precipitation time series of 12 GCMs was evaluated and validated against observed extreme precipitation in the time interval 1981–2010, which is the reference period of this study. Four statistical matrices (RR, MAE, PBIAS, RMSE) were used to measure the accuracy of the annual and seasonal maximum precipitation time series of 12 GCMs in reproducing the daily seasonal and annual maximum time series of observed precipitation. The correlation coefficient (RR), mean average error (MAE), percentage of bias (PBIAS), and maximum precipitation weighting root mean square error (PWRMSE) performance measure criterion were used to evaluate the 12 GCMs output with respect to the observed annual maximum precipitation. The results obtained for five locations in Awash basin are given in Fig. 3. Overall, the selected stations showed lower medium, and MAE, PWRMSE, and PBIAS values, respectively, indicating that the reliability and accuracy of the GCMs precipitation in reproducing observed seasonal and annual precipitation is high. The statistical matrix values were found different for different climate models. The MAE values vary in the ranges of 0.1–20 for annual maximum precipitation, PBIAS values range –40 to 30, PWRMSE values range 0–0.1, and RR –0.25 to 0.55 in all the selected catchments. Relatively, the bias correction methods performed better in C1 and C3.

The climate model biases must correct before using it to force the hydrological models, especially because of the nonlinear response of the hydrological model to the precipitation and temperature forcing mechanisms. Figure 4 shows the comparison of three bias

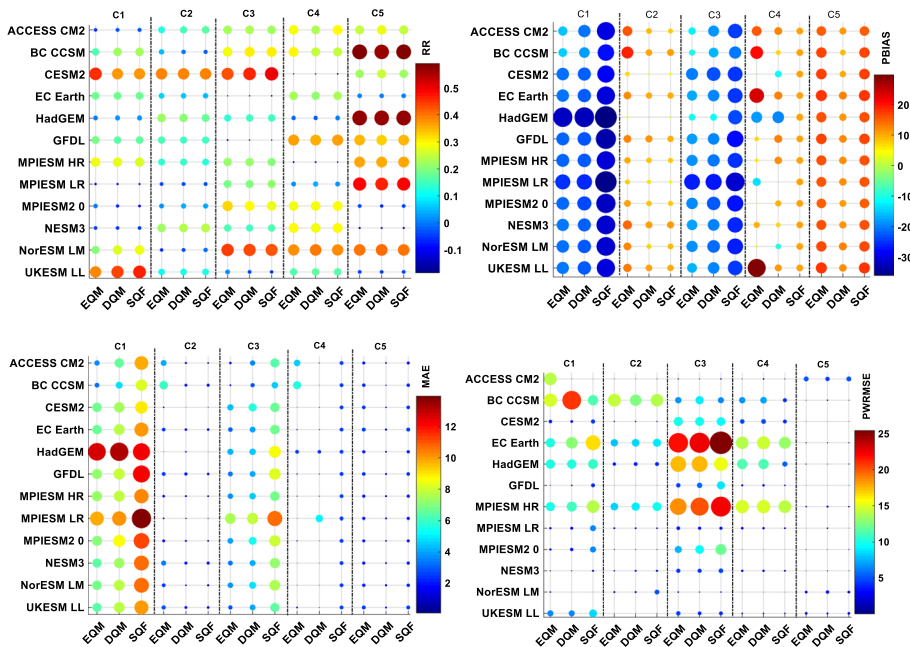


Fig. 3 Comparison of three bias correction techniques with observed daily maximum precipitation using a statistical matrix. The column represents each catchment using three bias correction methods, and each row represents each climate model

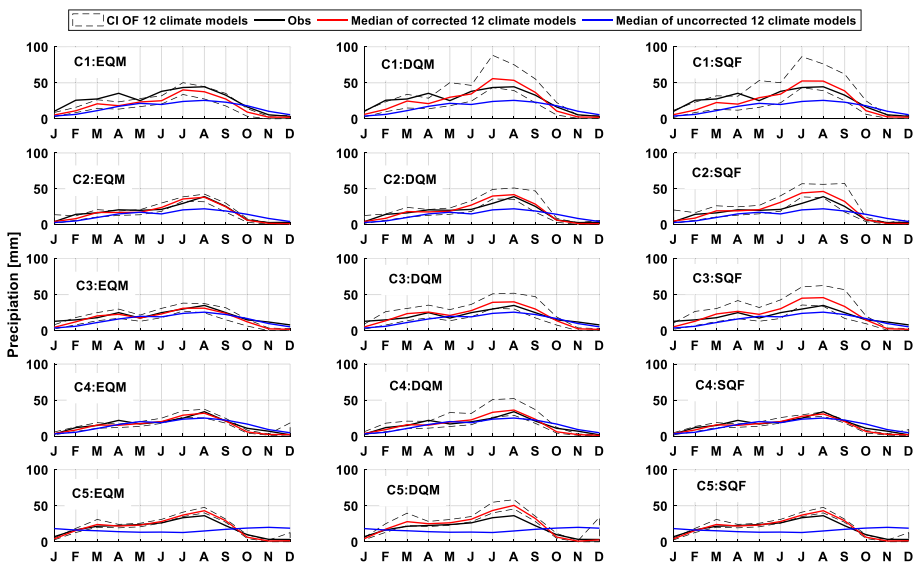


Fig. 4 Comparison of seasonal raw and bias corrected daily maximum precipitation simulations from 12 CMIP6 GCMs for each of our five study catchments. Each column presents results of one bias correction method of five catchments. Each row presents a result of three bias correction methods of one catchment. In each panel, the dashed black line is the range of bias-corrected 12 GCMs models, red line is the median of the corrected 12 GCMs, blue line is the raw median of 12 GCMs, and black line represents the observed monthly maximum precipitation

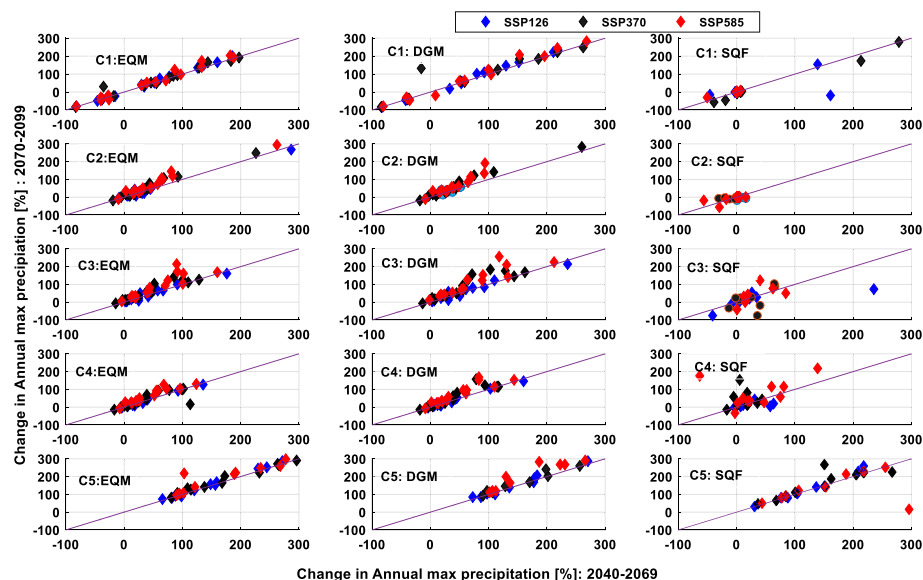


Fig. 5 Maximum precipitation change using three bias correction techniques and three scenarios. SQF-change factor, DGM-quantile mapping using gamma distribution, EQ- empirical quantile using simple interpolation. The blue dot stands for SSP126, red for SSP370, and green SSP585

correction methods on simulation of monthly maximum precipitation average over each catchment for the CMIP6 ensemble (color bands, single models shown as a straight line), and the observed monthly maximum precipitation is represented by blue straight-line color for the period 1981–2010. Overall, the 12 GCMs give a broader spread in the rainy season and a relatively narrow band of climate models in the dry season. The spread of the selected 12 climate model ensemble does not exceed these climate models' upper and lower limit in the reference period. However, the width of these 12 climate models mainly depends on the type of bias correction method. DQM and SQF methods are relatively given a narrower band/spread than the other, whereas the EQM provided a wide range of climate model spreads (Fig. 4). In general, the annual cycle of maximum precipitation shows that the uncorrected GCMs have a considerable bias with respect to observed time series. The result of the three bias correction techniques is not the same in reproducing the observed seasonal maximum precipitation of the selected catchments (Fig. 4). The performance of these bias correction techniques in the reference period (1981–2010) is uniform in five catchments. Using distribution quantile mapping and empirical interpolation, the corrected precipitation range/band of 12 climate models is relatively smaller.

These climate bias correction methods were applied to each GCM in future period (2015–2100). The influence of each climate bias correction method on the magnitude of change was assessed (Fig. 5). Daily precipitation between 2040 and 2069 (future-clim2) and 2070–2099 (future-clim3) was compared to the reference period 1981–2010 (Fig. 5). An increase was simulated in the far and near-future period by all GCMs scenarios in terms of annual maximum precipitation. The individual climate models simulated both increase and decrease in annual maximum precipitation. Similarly, changes from each bias correction approach show a slight difference in magnitude and direction in changing annual precipitation. Generally, these climate change signals are higher for

Table 2 Overall mean (Mean temp [°C]) and maximum daily temperature (Mean temp [°C]) conditions for the reference and the SSP126, SSP370 and SSP585 scenarios in 2050 and 2080 time slices

	Station	Reference	2050 conditions			2080 conditions		
			SSP126	SSP370	SSP585	SSP126	SSP370	SSP585
Mean temp [°C]	C1	13.37	14.25	21.05	23.55	16	23	28
Max temp [°C]		25.58	26.3	33.15	35.85	30	37	42
Mean temp [°C]	C2	13.02	13.75	20.55	26.05	15	22	27
Max temp [°C]		25.19	25.8	32.65	35.35	29	36	41
Mean temp [°C]	C3	12.91	14.75	19.05	23.05	17	19	27
Max temp [°C]		25.06	26.8	31.15	35.35	31	33	41
Mean temp [°C]	C4	12.57	14.25	21.05	20.05	16	23	21
Max temp [°C]		24.68	26.3	33.15	32.35	30	37	35
Mean temp [°C]	C5	13.14	13.75	20.55	23.55	15	22	28
Max temp [°C]		25.32	25.8	32.65	35.85	29	36	42

SSP585 than SSP370 and SSP126 and slightly higher in the C4 and C2 catchments. Generally, the changes are smaller in C1 and C4 using DG methods, whereas C2 and C3 changes are smaller using EQM and DQM. Interestingly, a linear relationship exists between the annual maximum precipitation changes in 2020s and 2050s with proportional magnitude. This indicates that the slope of these changes in 2020s, 2050s, and 2080s is a positive trend.

Three scenarios have been assembled from 12 CMIP6 climate models (Table S1), constructed for three SSPs. Figure S1 and Table 2 show the projected max air temperature under each SSP scenario for each of the 12 climate models, together with the mean of the ensemble air temperature projected for each SSP. Air temperature is more stable and consistent with time and space in Ethiopia. The projected air temperature from the new climate dataset and new scenarios provides reasonable temperature change, increasing by 1.7–3.6 °C and relatively uniform in the selected catchments (Fig. S1 and Table 2). However, the spreads of the selected climate models are more extensive in C1 and C5. Overall, the projected air temperature has a positive slope and continuously increasing in the clim1, clim2, and clim3 future period. Overall, the mean daily air temperature and maximum daily air temperature over the baseline condition for the three different climate scenarios and two future periods (2050's and 2080's) are presented in Table 2. It can be seen that the maximum and mean daily air temperature increase from the lower to the higher/worst climate scenarios in 2050's and 2080's periods.

4.2 Calibration of the Hydrological models and parameter uncertainty evaluation

An overview of hydrological model calibration performance was provided to assess the quality of the sets of robust reference parameters and to check and control that the models perform reasonably well in validation and calibration period. The calibration performance of the four hydrological model in terms of goodness-of-fit criteria (GOFC) is summarized as boxplots according to models and criteria (Figs. 6 and 7). The horizontal line (red lines) in the box is the median (0.5) of goodness-of-fit values, and the lower (0.25) and upper (0.75) envelopes show the 25th and the 75th percentile values and the lower and upper whiskers show the 5th and 95th percentile values, respectively.

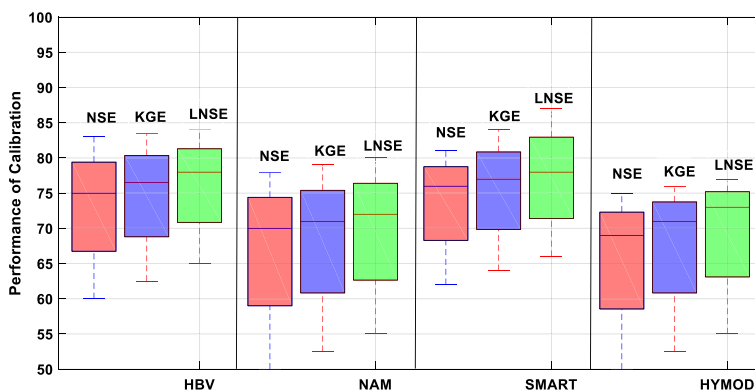


Fig. 6 Summary of GOF values across the five catchments for the calibration of the four rainfall-runoff models for the modeling periods according to the goodness-of-fit criteria

An overall analysis of the performance of the models showed that the SMART and HBV models are more efficient than the NAM and HYMOD models. The plots' median and upper whisker are lower for HYMOD models than for the NAM, SMART and HBV models. The number of runs with at least 60 for GOF is higher for SMART models than for HBV, NAM and HYMOD models. The concept of SMART is more applicable to the study catchments investigated, which is probably due to the structure of SMART models, which have more routing reservoir instead of none for HYMOD models. The SMART models are slightly more effective for performance than the HBV model in the given period. This is also confirmed by the number of runs, which is higher for HBV models than for the NAM model.

The results of model performance are more distinct for LNSE than for the two other GOFs because of the lower whisker and median, which is higher than for others (NSE and KGE). In the case of KGE, the upper whisker is slightly similar to the two other GOFs. The different versions of the HBV and HYMOD models with two or three parameters are almost equivalent in calibration period, despite the different model parameters to calibrate the performance criteria. It seems that assimilating the soil reservoir parameter to a water holding capacity of the soil is appropriate. It must be due to the fact that the two models are not very sensitive to this parameter.

The HBV, SMART, NAM and HYMOD hydrological parameter sets with a sample size of 30,000 were generated through uniform distributions. The 30,000 results from the Monte Carlo simulation are flagged with the NSE threshold criteria and carried forward to estimate future streamflow. GLUE-based parameter uncertainty approach was adopted to simulate a possible ensemble of daily streamflow, and NSE objective function was applied to separate the behavior and non-behavior simulations. In this study, the runoff simulation ensemble with likelihood value larger than 0.5 NSE threshold value was selected. The likelihood value less than the NSE threshold value of the model parameter sets is considered as non-behavioral. Figure 7 shows the seasonal maximum flow simulation results from best parameter (behavioral) sets, including observations in 1981–2010 period and confidence interval of 5% and 95% of the streamflow simulations. It indicated that the maximum seasonal simulation falls mostly between the 5% and 95% confidence bands, except in the C4 catchment. Almost in all the catchments, the upper 95% confidence interval band is mostly exposed to extreme seasonal peak flows. Overall, 83% of the observed maximum flow time

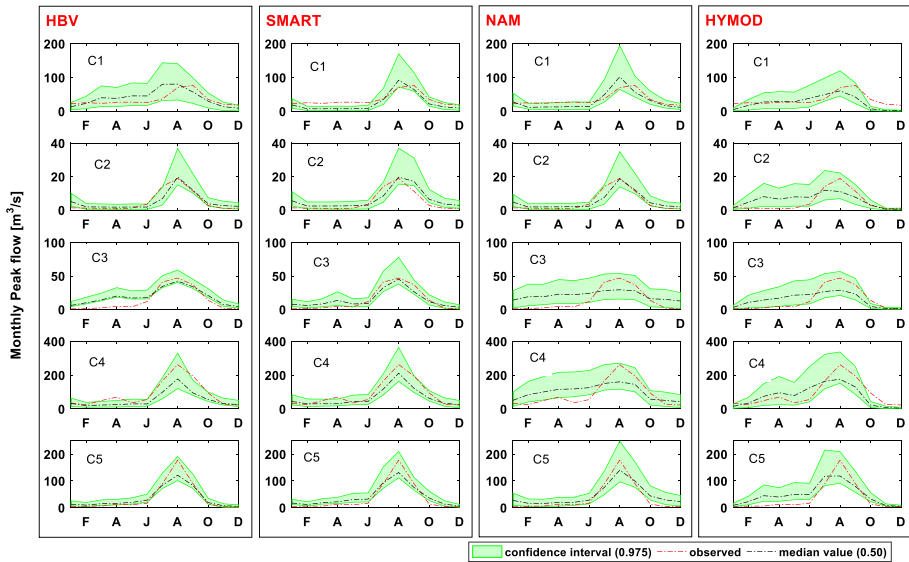


Fig. 7 The multi-model ensemble of monthly maximum flow simulation using best HBV (first column), SMART(second column) NAM (third column) and HYMOD(fourth column) parameter sets selected based on best NSE value range in the reference period at each catchment

series falls under the shaded area confidence interval and has a wide range of NSE values (0.3–0.75), and it is promising to use for climate change impact study.

4.3 Projected hydrological flow: performance of hydrological models under different bias correction approach

The long-term internal variability and projected flow dynamics from the period 1981–2100 at the given five stations are shown in Fig. S2. The x-axis is time in the year extending out to 2100, and the y-axis is the seasonal value of projected flow. It is clearly seen that the main rainy/peak flow season is covered by a yellow color, which is mainly June, July, and August. The flow during flood seasons has increased noticeably for each station since 2010. Relatively, there is a strong increasing trend in seasonal and annual flow in C1 than in C2. Overall, considering the daily projected flow time series shows that the peak flow increased for the coming three periods (2010–2039, 2040–69, and 2070–2099). This increase appears to be more significant during the summer and early spring season, with a mean peak flow increasing from 60 to 70 m³/s and during summer, with a mean flow increasing from 85 to 100 m³/s at C3. Also, projected flows seem to be smoothed during late autumn and early summer periods, with a mean flow of about 35 m³/s. In contrast, the future conditions show an inevitable variability from about 25 m³/s in May to 55 m³/s in October (Fig. S2).

The changes in flow simulations forced using three bias correction procedures for climate variables have been compared and presented in Fig. 8. Overall, it observed positive changes in annual maximum flow using DQM, EQM, and SQF, while the magnitude of annual maximum flow changes is not the same. However, these are not uniform across the selected catchments, bias correction methods, and climate scenarios. Relatively, distribution-based bias correction models give smaller changes in the annual maximum flow time

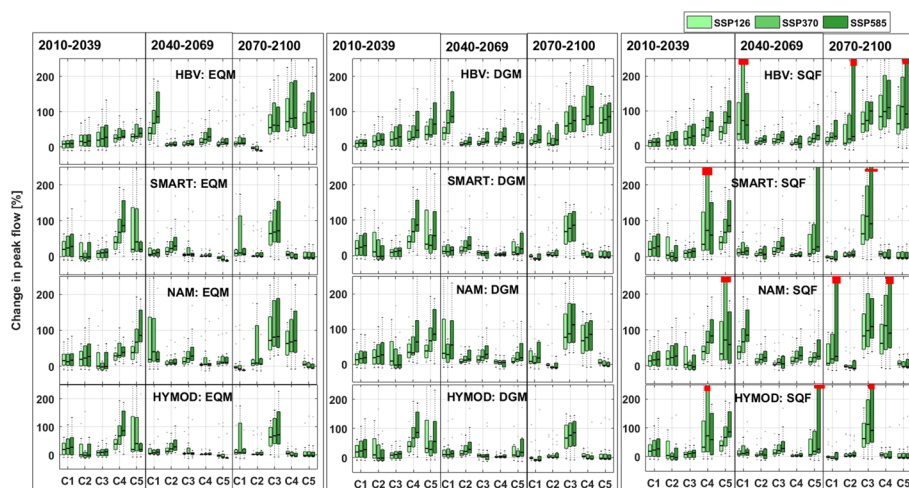


Fig. 8 River flow change in clim1 (2010–2039), clim2 (2040–2069), and clim3 (2071–2100) with respect to the reference period (1981–2010). Each box plot represents the spread of twelve climate models listed in Table S1 with median 0.5 and 0.25 and 0.75 quantiles. Each column stands for each scenario (the first column for SSP126, the second column for SSP370, and the third column for SSP585). Each column represents each catchment (first column C1, second column C2, third column C3, fourth column C4, and the last column for C5). The red box at the top of the box plot in the right side indicates extreme outlier points

series. This indicates that wet day frequency correction is significant in understanding the future maximum flow projections. Mainly, annual maximum flow changes using EQM give a higher spread range and uncertainty. Boldly observed that annual maximum flow changes have smaller uncertainty (size of the box plot) in SSP370 and higher in SSP126 (Fig. 8). Therefore, correcting the wet days and intensity of precipitation may significantly change the change magnitude and direction and minimize uncertainty. The annual maximum flow changes are smaller in the C1 and C3 catchments and higher in C5 catchment.

4.4 Flood hazard projections under varying climate conditions

Best fitted to annual maximum flow was selected for each station and climate model, and the best distribution was selected based on Akaike information criteria (AIC) (Table 3). Across the selected catchments and climate models, GEV, Gamma, Pearson type III, and Weibull distribution are the most dominant distribution types (Table 3). The projected annual maximum flow series has different distribution characteristics, resulting from different flood hazard magnitude and risk level. Of which, the three most dominant and best-fitted distribution models have similar probability density functions (PDF).

Using the most dominant PDF model (GEV), the flood quantile changes due to climate change have been evaluated. Figure 9 shows flood quantile changes in climate1 clim2 clim3 with respect to the reference period, estimated using an ensemble of 12 GCMs and three bias correction methods. The ensemble average of 12 climate models within three scenarios assessed (SSP126, SSP370, and SSP585) and three bias correction methods in five catchments shows a significant increase in flood quantile magnitudes and frequency in future (Fig. 9). The changes of flood quantiles are mostly consistent with changes in maxima flow, but it does not mean that the magnitude of changes and directions are the

Table 3 Best fitted frequency distribution models to simulated peak flow using 12 GCMs and four hydrological models (SMART, NAM, HBV and HYMOD). Each frequency model is represented by number: Pearson-III distribution by 1, GEV by 2, Gamma by 3, Weibull by 4, Nakagami by 5, lognormal by 6, Loglogistic by 7 and Inverse Gaussian by 8

	SMART					HBV					NAM					HYMOD				
	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
Observed	1	2	2	3	5	3	3	2	3	4	2	2	3	4	4	3	2	3	3	3
ACCESS-CM2	1	2	2	2	1	1	2	2	2	1	2	2	2	1	1	1	2	2	2	2
BCC-CSM	2	5	1	7	8	5	5	3	5	7	3	4	5	2	6	4	4	1	1	5
CESM2	2	5	5	2	4	3	4	5	3	4	4	4	4	3	4	8	4	4	4	4
EC-Earth	6	1	7	8	2	4	3	5	2	4	2	5	1	5	3	4	5	5	5	5
GFDL	2	2	6	2	1	2	2	4	3	2	3	3	3	2	1	2	3	2	3	2
HadGEM	3	5	2	2	2	3	4	3	2	2	8	3	2	2	3	3	2	3	3	3
MICRO6	5	6	1	6	2	4	5	3	1	3	4	2	3	4	3	4	4	3	3	4
MPIESM_HR	4	2	4	7	4	4	3	4	5	5	3	4	5	5	4	1	4	4	4	3
MPIESM2	4	1	7	8	3	4	2	2	6	5	1	5	6	1	3	3	4	5	2	5
NESM3	1	7	5	4	6	4	5	5	5	5	4	1	5	5	5	5	5	5	5	5
NorESM_LM	2	2	3	3	2	2	2	3	3	2	2	3	3	2	2	2	2	3	2	2
UKESM_LL	8	5	2	4	6	7	6	4	4	5	5	4	4	1	6	6	5	4	5	5

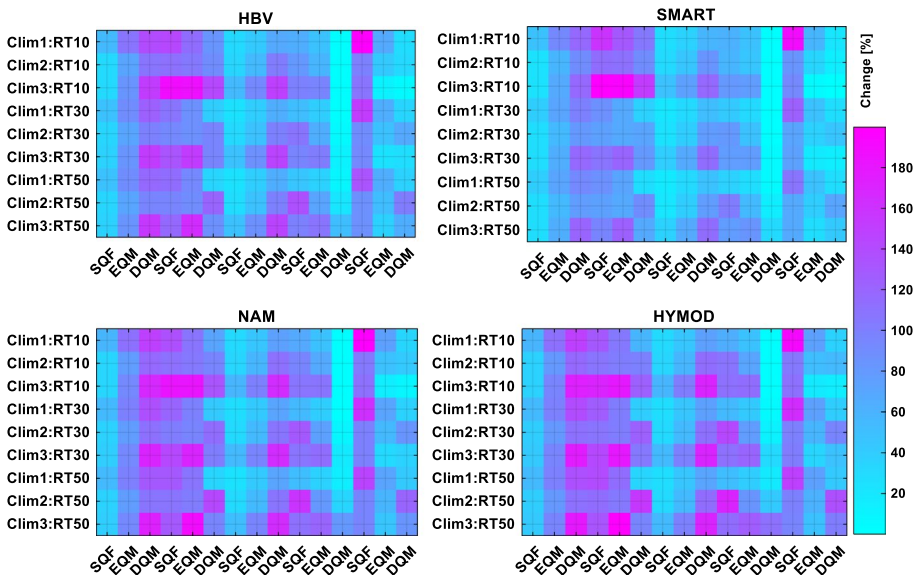


Fig. 9 Changes in peak flow quantiles using three bias correction techniques and GEV frequency distribution model (at 10-, 30- and 50-years return period) using four hydrological model (HBV, SMART, NAM and HYMOD) at each catchment

same. The future flood quantile changes are not the same in space and bias correction methods, and the frequency of the larger return period flood increases to higher than once in 20 years. The smallest changes are observed in C1 using EQM, C4 using SQF, C5 using DQM, and C3 using EQM. Relatively, the methods are consistent in C2 and C3 in providing information about flood quantile changes in future. Simultaneously, the flood quantile

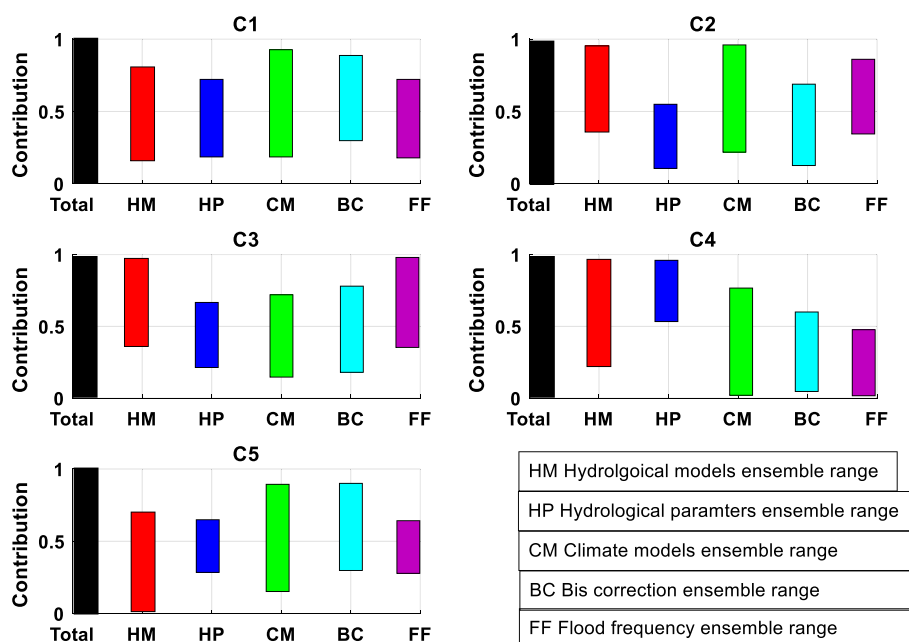


Fig. 10 Summary of the additive way of uncertainty contribution analysis result using the direct additive method without considering their interaction. The peak flood was estimated at 30 years return period and calculated its change with respect to far future period

changes are significant at different return periods with different hydroclimatic simulations. For example, C1 catchment has a change range from 0 to 100%, but using the DQM, the changes range from 0 to 170%. Similarly, in C5 catchment 0–100% using DQM.

4.5 Uncertainty estimation and decomposition of associated sources in flood hazard estimation

The four sources of uncertainty were estimated based on the one chain principle. The one chain uncertainty band quantification started from one climate model that passed through three bias correction techniques and four hydrological models with best hydrological parameters (Fig. 10). The contribution of each source of uncertainty is not uniform across the selected catchments. Each color band represents each source of uncertainty as a relative range of cascade uncertainty. Spreads of the 12 climate models represent the climate models uncertainty band; the spread of four hydrological models for hydrological structure uncertainty; the hydrological parameter uncertainty is presented by 95% confidence interval of selected hydrological parameter sets. The uncertainty due to bias correction types is estimated from the spread of three bias correction simulation results are estimated. However, the ensemble uncertainty of GCMs is relatively more considerable than the other sources. By comparison, hydrological parameter sets do not play a significant role in future hazard estimation (Fig. 10). Future flood hazard in C1 catchment is highly sensitive to climate change and respective bias correction methods, while C2 is more sensitive to hydrological model structures and bias correction methods in flood hazard estimation. Overall,

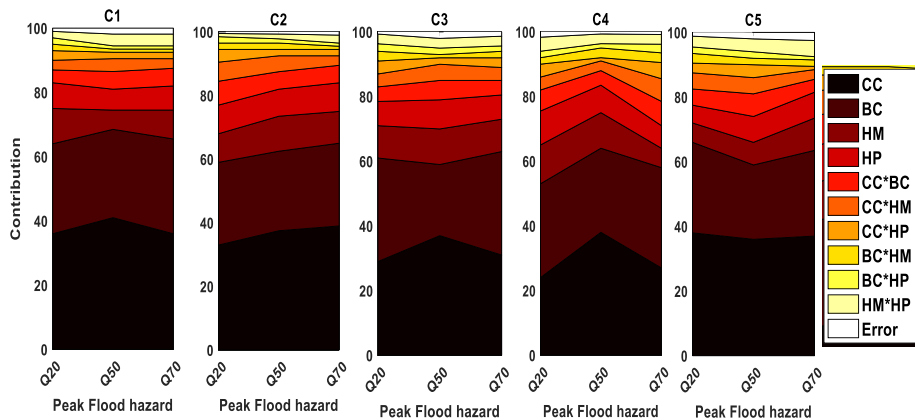


Fig. 11 Contribution of each source of uncertainty in flood frequency estimation under climate change for the five catchments identified as C1 to C5. The three Quantiles estimated at return periods (RT) of 20, 50 and 70 years (Q20, Q50, Q70) arranged in the x-axis from left to right, respectively. Each color represents each source of uncertainty: green is uncertainty from climate models (CC), red from bias correction methods (BC), orange color from hydrological model structure (HM), pink color from the hydrologic model parameter (HP), and the other colors represent the interaction between the main factors

from the highest to lowest the sources are arranged as follows: the climate models, hydrological model structure, bias correction procedures and hydrological parameters.

Figure 11 shows only the main sources of uncertainty and their respective bands at specific return periods and catchments, with considering the main factors' interdependence using ANOVA (Meresa and Gatachew 2019), the contribution of individual sources of uncertainty as the main source and their interaction contribution to the flood hazard magnitude and frequency changes in future. The decomposition of source uncertainty from the main variables and their interaction were separated based on the variance in flood hazard values and calculated out of 100, which means the sum of all sources is 100. Hydrological parameter sets on flood hazard projection are not significant and less uncertain in the selected catchments. This finding is more related to (Meresa et al. 2021) conclusion. In comparison, the climate change variability shows a significant impact and uncertainty on the projection of flood hazards. This may be due to the projected precipitation intensity, higher temperature, and time of concentration in the area. On the other side, the contribution of hydrological parameter sets is not significant in flood hazard estimation due to weak role of ground and interflow components in the peak flow (Fig. 11). In general, the climate models and hydrological models structure are the main factors in peak flood magnitude and frequency projection, Whilst the uncertainty from the hydrological parameter sets is not added significant value to the total uncertainty. The share of climate models, bias correction, and frequency distributions are significant at design flow value estimation at 70 year return period. In C1, future flood hazard risk is highly uncertain due to climate model variability and their respective bias correction methods. By contrast, C2 catchment is more sensitive to the frequency distribution and climate model variability.

5 Discussion

5.1 Projected hydroclimate using CMIP6 dataset

This study used CMIP6 climate change dataset source to understand the impact and uncertainty in future peak flood hazards. From CMIP6 GCMs, 12 GCMs climate models with three climate scenarios (SSP126, SSP370, SSP585) were applied for future hydrological extreme simulations. Due to the lack of robust atmospheric-land model interaction, uncorrected GCMs are not recommended for extreme hydrological modeling (Meresa et al. 2017; Mendez et al. 2020). However, the impact of climate change magnitude also depends on the bias correction methods' principles (Pierce et al. 2015). Three climate bias correction methods were applied to correct the GCMs output with daily observed air temperature and precipitation.

Overall, the distribution-based bias-corrected model fields successfully reproduced the observed maximum precipitation, both in terms of the number of wet days and precipitation intensity. This implies that correction of wet days and intensity of GCMs is significant and provides a significant role in minimizing uncertainty, whereas the change factor and empirical methods give similar results in reproducing the observed daily mean air temperature. This shows that the interval variability and magnitude of the climate variable significantly influences the result of each bias-correction procedures. Due to this reason the bias correction to the GCM temperature field was small compared with precipitation.

Moreover, we found that the maximum future temperatures will increase but precipitation may either increase or decrease depending on climate model, climate scenario, hydrological model, bias correction method, future climate period and catchment. At each stage and modeling chain, the peak flood magnitude and frequency gives different change magnitude. The highest change in annual maximum precipitation was observed in C4 catchment, whereas in C5 the changes in the clim2 and clima2 show a non-significant (near to zero) magnitude and direction. In C1 catchment, most climate models indicate a positive change, but these changes are very small. These changes in precipitation and air temperature are consistent with previous findings using RCP-AFRO-CORDEX (Meresa and Gatachew 2019; Taye et al. 2015), using MIROC-ESM-CHEM and CSIRO-Mk3-6-0 (Daba et al. 2020), and using Statistical Down-Scaling Model, SDSM-CMI5 (Gebrechorkos et al. 2020) in Ethiopia. The annual daily maximum precipitation and daily mean air temperature projected under SSP126, SSP370, SSP585 shared socioeconomic pathways scenarios suggests that wet conditions could be slightly intensified and increased in C1 and C5. However, it doesn't mean that the mean peak flood is much higher the normal condition. In contrast, the magnitude of the positive changes in maximum precipitation will be intensified in C1 and C2. Similarly, the mean temperature changes of these selected catchments are not much different in magnitude and direction. However, SSP126 gives lower changes in precipitation and temperature magnitude than changes using SSP370 and SSP585.

5.2 Modeling and projections of hydrological extremes

We examined the impact of the choice of different hydrological model structures with respect to reproducing historical extreme flows and the effect of the model structure choice on hydrological climate change evaluations. We studied four hydrological models with different number of parameter and routing and storages. The hydrological model with three storages and five parameters and six storages with 12 parameters yield equally good results

in predicting peak flow in the selected five catchments. Those different model structures reproduce historical flow well and the projections of hydrological response under changing climate are consistent with meteorological data input patterns.

We compared the feasibility and robustness of four conceptual models widely used in Africa and Europe: HBV, SMART, NAM and HYMOD models. The investigation was undertaken on five catchments covering the main Awash River basin. The model's performance were evaluated under three GOF: KGE, NSE and LogNSE. For the climate change impact study, only those obtained with parameters sets giving at least 55% (0.55) as the calibration and validation performance were considered. An analysis of validation and calibration performance showed that SMART model performs better than HBV model, regardless the GOF. The performance of NAM models is somewhat similar, albeit slightly higher than that of HYMOD models. The behavior of and validation and calibration performance of these models differs according to the GOF criterion considered and is different from model to model. A clear relationship was found between validation and calibration performance and the hydrological flow regime using four hydrological models and three objective functions.

Robustness analysis also showed SMART models to be more robust than NAM and HYMOD models. The hydrologic models are most robust in terms of LogNSE. They seem to perform and work better when the same weight is given to all parts of the hydrograph. HBV and SMART models are the most robust in terms of NSE for peak flood simulation. In terms of KGE, the HBV model is the most robust and HYMOD model shows lower KGE values. SMART is the most robust in terms of LogNSE than any other likelihood functions. The analysis of the robustness and uncertainty of the models structure according to input climatic variations between validation and calibration periods (precipitation and PET) showed that it is better to transfer the model parameters from a higher PET and a drier period than doing the reverse. A more refined analysis showed that a precipitation variation of -9% to 9% or PET of -4% to 4% between the calibration and validation period causes: a loss of robustness less than 10% for SMART models and 30% for HBV models in terms of NSE and KGE; a loss of robustness between 15% and 20% for SMART models in terms of LogNSE. Authors such as Coron et al. (2012) and Dakhlaoui et al. (2017), reached conclusions comparable to those in this study. Particularly, Coron et al. (2012) found that the difference among the hydrological models and between validation and calibration periods should not exceed 15% for better model parameter transfer results. The sensitivity of the model parameters depends on the hydrological processes that predominate in the river flow over the calibration period and hydrological model structure. Similarly, the more sensitive the model parameter over a given period, the more likely it is to be the optimal model parameter for other periods.

Five catchments have been selected from different hydroclimatic and geographical locations to investigate the impact of climate change on flood hazard and identify the associated uncertainty in the projected flood hazard magnitude and frequency. Three hydrological model was calibrated and validated against the observed streamflow data using Nash–Sutcliffe (NSE), LogNSE and KGE objective functions in the period of 1981–1990 and 1991–2005, respectively. A likelihood value higher than 0.5 indicates a satisfactory hydrological simulation zone (indicates how well the hydrological model reproduces the extreme high flow) (Bennis and Crobeddu 2007) and promising for further analysis. Therefore, 0.5 was selected as a threshold in runoff simulation and accepted for future runoff simulations. The five catchments have an NSE value ranging from 0.5 to 0.75, LogNSE ranging from 0.5 to 0.78, and KGE ranging from 0.5 to 0.71 from the 30,000 parameter sets sample. Similarly, other studies also show similar results using HEC-HMS (Roth et al.

2018), using SWAT model (Shawul et al. 2019), and using the HBV light model (Bekele et al. 2019) in Awash catchments.

Uncorrected and the future corrected precipitation and temperature data from CMIP6 data sets were used to force four hydrological model and to examine the potential impact of climate change on the future flood hazard/peak flow. Overall, the projected high flow changes are positive, which is an increase in high flow in the near future (2010–2039), far future (2040–2069), very far future (2070–2099) for an ensemble of SSP370. Like the precipitation changes, the effect of climate bias correction on high flow change is clearly visible and brought higher changes in magnitude with a large uncertainty band. However, the distribution-based quantile mapping techniques relatively produce reasonable changes in high flow at the selected sites. The changes are higher using SSP126 and SSP585 climate scenarios.

In contrast, the SSP370 gives a smaller band in high flow changes in a near and far future period, which is a direct reflection of maximum precipitation change. In the C1 and C4 catchments, the near future changes, far future, very far future for an ensemble of SSP370 shows significant changes with a large uncertainty band, whereas in C2 and C3 catchments, the changes are not uniform across the bias correction methods, and changes are smaller than C1. This is most likely due to increased summer and spring precipitation (Meresa and Gatachew 2019). Similarly, the distribution-based climate bias correction gives relatively smaller comparisons with the empirical and linear interpolation methods at most of the catchments.

5.3 Flood hazard frequency and magnitude estimation under climate change

The impact of climate change on flood hazards was assessed through the extreme frequency distribution models. Each catchment's baseline and future flood frequency simulations were estimated from 12 climate models (GCMs). Various distribution types are commonly used in the US, Europe, and Africa (Kay et al. 2009; Collet et al. 2017; Meresa and Romanowicz 2017; Meresa and Gatachew 2019; Lawrence 2020). In this study, we fitted a series of frequency distribution models to the annual maximum flow time series and assessed the best performing distribution model. The GEV distribution model is commonly used in hydrological extremes analysis and was found to be the best statistical description of the model prediction data in this study. It was applied for flood quantile estimation and associated impacts. The estimated magnitude of flood quantiles at different return periods using GEV frequency distributions is not the same in magnitude and uncertainty band. Each climate model at each catchment gives a very wide range of flood quantile values. Overall, smaller flood quantiles were observed in C3 catchment, whereas the highest was estimated from C1 and C4 catchments. The estimated quantiles' total range is relatively smaller in C3 and C2, whereas the highest uncertainty range is estimated on C4 and C1 catchments.

Moreover, the changes of flood quantile values are not the same (significance difference) across the catchment and type of bias correction methods. In the C2 and C3 catchment, the flood quantiles at different return periods are expected to be larger in the near, far, very far future. In comparison, the changes are relatively higher in C1 and C4 catchments and expected to be higher in the very far period. This is mainly due to the magnitude and distribution of flood events in the given period (Armstrong et al. 2012; Rao et al. 1972; Keast and Ellison 2013).

5.4 Uncertainty in flood hazard frequency and magnitude and decomposition

In this study, 12 climate models, three bias correction methods, three hydrological models with 30,000 parameter sets, GEV flood frequency model, and three climate scenarios from the new dataset (CMIP6) were considered to analyze the impact of future flood hazard in the selected five Awash catchments. Using additive uncertainty aggregation, each component of these analyses is integrated into one framework that contains information from each part. Climate models are highly uncertain in characterizing the future climate variables. In the last decades, various studies stated that RCP and SRES data are very weak in reproducing the historical climate extremes (Fowler et al. 2007; Kingston et al. 2011; Saini et al. 2015; Meresa and Romanowicz 2017). This uncertainty is may be due to the structure, parametrization, and spatial resolution of the GCMs. Using multiple models in climate change impact analysis would lead to an improved understanding of the uncertainty associated with climate models. It is essential in flood risk magnate and water resource management. Similarly, in this study, multiple climate models from the new data set CMIP6 (shared socio-economic pathway scenarios) were evaluated for flood hazard projection in Awash basin. The spread/discrepancy of climate models' impact is significant. There are also important uncertainties component which is associated with the climate bias correction methods. The catchment-scale characteristics are not the same as the GCMs spatial and temporal characteristics; therefore, it is important to use highly relevant climate bias correction techniques to understand the uncertainty and minimize. Three bias correction methods were used to correct the GCMs output at catchment scale. These methods were effectively corrected in the historical period and then adopt the parameters for future climate corrections. There are few studies related to bias correction methods uncertainty and flood frequency (Kay et al. 2009; Saini et al. 2015; Soriano et al. 2019). They found that the climate bias correction could alter the magnitude of flood hazard. This study also confirmed this study using three climate bias correction methods and GCMs from CMIP6 in Ethiopia's selected catchment. We find that the uncertainty source due to wet days and intensities of precipitation has a significant impact on the flood hazard magnitude (e.g., C4 catchment).

The hydrological model structure and parameters represent hydrological process. The hydrological parameters control the flow from single droplet of rain to the deep catchment percolation. In this study, more sensitive parameter sets of HBV, SMART, NAM, HYMOD hydrological models were selected using NSE objective function values (0.5 as threshold). GLUE was used to investigate the role of hydrological parameters in flood hazard estimation in the selected catchments. The analysis was performed by using a uniform sampling approach from the given parameter ranges. The simulated flow in the historical period is pretty good for mean and high flow simulation. The width of the simulated band is not significantly increased by changing the threshold for high flow simulation. Generally, the uncertainty is band due to parameter change is not significant in flood hazard estimation. Our finds conformed that the role of hydrological parameter sets is less in all the selected catchments in Awash basin. Yan et al. (2015), Meresa and Romanowicz (2017), Joseph et al. (2018) they appear with similar conclusion that high flow is less sensitive to hydrological parameters.

The uncertainty related to flood frequency under climate change is not extensively investigated and it has been very challenging to integrate with other sources of uncertainties. Each extreme high flow simulation from each climate model simulation and bias corrected simulations were fitted to most dominant frequency models. Several researchers

have been conducted flood frequency estimation, and they concluded that GEV and EV distribution models are the most commonly used frequency models in Awash catchments (Tegegne et al. 2020; Ahilan et al. 2012). However, our numerical experiment result confirms that GEV distribution is the most dominant model type in the selected Awash catchments. The uncertainty band is not significant in the selected catchments of flood quantile values. Therefore, the uncertainty of flood frequency models in these catchments are avoidably in the hydro-climate projections. However, the results should be assessed with care because the probability of recurrence of extreme floods of the selected catchments are become more intense and frequent and intense in the near, far, and very far future periods.

6 Conclusions

The results of climate variables from CMIP6 data set are consistent with previous studies and applied for the first time in estimation of future flood hazard under varying climate conditions and associated uncertainty estimation. There is a significant range of differences among the results of GCMs in climate projections. The results also confirm that the CMIP6 GCMs outputs are promising for climate change impact studies on hydrological extremes. Our results also indicate that the proper climate model bias correction techniques can improve future estimated precipitation and temperature characteristics. The climate models bias has been corrected using three techniques (DQM, EQM, and SQF). The precipitation results using the DQM approach are relatively close to observed daily maximum precipitation data. This indicates that climate models have higher biases in projection of extreme precipitation magnitude (greater than 75% of the maximum precipitation), and smaller bias results in lower precipitation magnitude (lower than 75% of the maximum precipitation). This is due to seasonal variability and intensity of precipitation. Therefore, giving more focus to the upper quantile of precipitation distribution can improve the model's capability to reproduce the observed time series and reduce the error of model predicted extreme values of precipitation and temperature.

The projected hydro-climate extreme indices are less consistent and indicate a strong model agreement with a wide range of variability in the selected catchments. However, the magnitude of changes in hydro-climate extreme indices mainly depend on the type of bias correction technique. DQM, EQM, and SQF give relatively smaller range of changes in annual maximum precipitation and annual maximum flow. Overall, increases in flood and extreme rainfall are expected and are not uniform in space/time. Similarly, the GEV frequency model was fitted to simulated annual maximum flow time series. The flood quantile changes were calculated based on the relative difference between the future (2015–2100) and historical (1981–2010) flood quantile simulations. Due to precipitation and temperature changes, in the twenty-first century, the magnitude of flood quantiles at different return periods between 5 and 100 years could increase between 5 and 50% in the selected catchments. However, these changes are not the same across all bias correction methods and frequency models in each catchment.

Comparing the results of climate change scenarios revealed that SMART, HBV, NAM and HYMOD are suitable for assessing climate change impact on peak flow because they provide similar and comparable results. The NAM and HYMOD models produce considerably higher peak discharge. The SMART model can simulate the slow and fast subsurface flow components and produce smaller change in peak flow in almost

all the catchments, all the catchments. Whilst the HBV model shows suitability for peak flow simulation in C1 and C4 catchments due to its physical representativeness and exhibits the best tradeoffs between calibration parameterization efforts. In contrast, the NAM model reproduces the peak flow curve very well but does not significantly contribute to understanding the subsurface processes. This finding strengthens the opinion that a smaller number of parameters and storages would lead to a poor representation of hydrological processes, especially in data-sparse catchments.

The uncertainty analysis was performed by comparing each source of uncertainty regarding the total aggregated value flood quantile at different return periods. The result confirms that hydrological model parameters, spread of climate models and bias correction techniques are important sources of uncertainty in flood hazard estimation. Specially, the climate models, bias correction methods and frequency distribution type models are the most significant source of uncertainty in future flood hazard estimation. This additive of uncertainty analysis is also proved by ANOVA analysis. ANOVA is variance-based sensitivity analysis and provides information about the main factors' contribution and their interactions. The decomposition of sources of uncertainty in future hazard estimation was performed based on the variance in changes of flood hazard value at Q20 (probability of recurrence of flood event at 20 years return period), Q50 (probability of recurrence of flood event at 50 years return period) and Q100 (probability of recurrence of flood event at 100 years return period). The result confirms that climate change is the dominant factor in C1 and C2, whereas C3 bias correction type and C4 distribution type are very important factors. The ANOVA also confirms that the contribution of their interactions is significant. The contribution of climate models and bias correction methods and interaction is especially substantial and non-negligible in flood hazard estimation.

Moreover, better flood risk management and policy need to consider these main factors and their management policy interaction. However, the value range of each source is large and challenging to communicate with decision-makers and stakeholders. Therefore, we strongly recommended considering the use of each spread's median values for future flood risk management and take as standard information on extreme flood adaptation measures.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11069-021-05152-3>.

Acknowledgements This study is supported by the DAAD and FIAT PANIS program. We extend our thanks to the Ethiopian Meteorology office and water resources minister for kindly providing the hydro-meteorological data of Awash basin.

Author contributions HM designed and performed the experiments, computational framework, derived the models, and analyzed the data. Also worked out almost all the technical details and performed the numerical calculations for the suggested experiment. HM, BT and TM contributed to the design and execution of the research, the analysis of the results, and the manuscript's writing. Finally reviewed, edited, and approved by BT.

Funding The funding was provided by DAAD and FIAT PANIS program (Grant No. 2018). Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors declare no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahilan S, O'Sullivan JJ, Bruen M (2012) Influences on flood frequency distributions in Irish river catchments. *Hydrol Earth Syst Sci* 16(4):1137–1150. <https://doi.org/10.5194/hess-16-1137-2012>
- Almazroui M (2019) Temperature changes over the CORDEX-MENA domain in the 21st century using CMIP5 data downscaled with RegCM4: a focus on the Arabian Peninsula. *Adv Meteorol* 2019:18. <https://doi.org/10.1155/2019/5395676>
- Armstrong WH, Collins MJ, Snyder NP (2012) Increased frequency of low-magnitude floods in New England. *J Am Water Resour As* 48(2):306–320. <https://doi.org/10.1111/j.1752-1688.2011.00613.x>
- Balke T, Nilsson C (2019) Increasing synchrony of annual river-flood peaks and growing season in Europe. *Geophys Res Lett* 46(17–18):10446–10453. <https://doi.org/10.1029/2019GL084612>
- Bastola S, Murphy C, Sweeney J (2011) The sensitivity of fluvial flood risk in Irish catchments to the range of IPCC AR4 climate change scenarios. *Sci Total Environ* 409(24):5403–5415. <https://doi.org/10.1016/j.scitotenv.2011.08.042>
- Beigi E, Tsai FTC, Singh VP, Kao SC (2019) Bayesian hierarchical model uncertainty quantification for future hydroclimate projections in Southern Hills-Gulf region, USA. *Water (Switzerland)* 11(2):268. <https://doi.org/10.3390/w11020268>
- Bekele D, Alamirew T, Kebede A, Zeleke G, Melesse AM (2019) Modeling climate change impact on the hydrology of Keleta watershed in the Awash River Basin, Ethiopia. *Environ Model Assess* 24:95–107
- Bennis S, Crobeddu E (2007) New runoff simulation model for small urban catchments. *J Hydrol Eng* 12(5):540–544. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:5\(540\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(540))
- Berg P, Feldmann H, Panitz HJ (2012) Bias correction of high resolution regional climate model data. *J Hydrol* 448–449:80–92. <https://doi.org/10.1016/j.jhydrol.2012.04.026>
- Bergström S (1976) Development and application of a conceptual runoff model for Scandinavian catchments. *Smhi, RHO* 7:134. <https://doi.org/10.1007/s11069-004-8891-3>
- Beven K (2007) Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrol Earth Syst Sci* 11(1):460–467. <https://doi.org/10.5194/hess-11-460-2007>
- Beven K, Binley A (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrol Process* 6(3):279–298. <https://doi.org/10.1002/hyp.3360060305>
- Beven K, Binley A (2014) GLUE: 20 years on. *Hydrol Process* 28(24):5897–5918. <https://doi.org/10.1002/hyp.10082>
- Blöschl G, Bierkens MFP, Chambel A, Cudennec C, Destouni G, Fiori A et al (2019) Twenty-three unsolved problems in hydrology (UPH)—a community perspective. *Hydrol Sci J* 64(10):1141–1158. <https://doi.org/10.1080/02626667.2019.1620507>
- Boyle DP (2000) Multicriteria calibration of hydrological models. PhD Dissertation, Dep of 5 Hydrol and Water Resour, Univ of Arizona, Tucson
- Charles S, Chiew F, Potter N, Zheng H, Fu G, Zhang L (2019) Impact of downscaled rainfall biases on projected runoff changes. *Hydrol Earth Syst Sci Discuss*. <https://doi.org/10.5194/hess-2019-375>
- Chen L, Asce SM, Singh VP, Asce F, Shenglian G, Hao Z, Li T (2012) Flood coincidence risk analysis using multivariate copula functions. *J Hydrol Eng* 17:742–755. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000504](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000504)
- Chen J, Brissette FP, Chaumont D, Braun M (2013) Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resour Res* 49(7):4187–4205. <https://doi.org/10.1002/wrcr.20331>
- Collet L, Beevers L, Prudhomme C (2017) Assessing the impact of climate change and extreme value uncertainty to extreme flows across Great Britain. *Water (Switzerland)* 9(2):1–16. <https://doi.org/10.3390/w9020103>

- Coulibaly T, Islam M, Managi S (2020) The impacts of climate change and natural disasters on agriculture in African countries. <https://doi.org/10.1007/s41885-019-00057-9>
- Daba MH, You S (2020) assessment of climate change impacts on river flow regimes in the upstream of Awash Basin, Ethiopia: based on IPCC fifth assessment report (AR5) climate change scenarios. *Hydrology* 7:98. <https://doi.org/10.3390/hydrology7040098> www.mdpi.com/journal/hydrology
- Dakhlaoui H, Ruelland D, Tramblay Y, Bargaoui Z (2017) Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *J Hydrol* 550:201–217. <https://doi.org/10.1016/j.jhydrol.2017.04.032>
- Ehret U, Zehe E, Wulfmeyer V, Warrach-Sagi K, Liebert J (2012) HESS opinions “should we apply bias correction to global and regional climate model data?” *Hydrol Earth Syst Sci* 16(9):3391–3404. <https://doi.org/10.5194/hess-16-3391-2012>
- Feng D, Beighley E (2020) Identifying uncertainties in hydrologic fluxes and seasonality from hydrologic model components for climate change impact assessments. *Hydrol Earth Syst Sci* 24:2253–2267
- Fowler HJ, Ekström M, Blenkinsop S, Smith AP (2007) Estimating change in extreme European precipitation using a multimodel ensemble. *J Geophys Res Atmos*. <https://doi.org/10.1029/2007JD008619>
- Gebrechorkos SH, Bernhofer C, Hülsmann S (2020) Science of the total environment climate change impact assessment on the hydrology of a large river basin in Ethiopia using a local-scale climate modelling approach. *Sci Total Environ* 742:140504. <https://doi.org/10.1016/j.scitotenv.2020.140504>
- Giorgi F, Gao XJ (2018) Regional earth system modeling: review and future directions. *Atmos Ocean Sci Lett* 11(2):189–197. <https://doi.org/10.1080/16742834.2018.1452520>
- Griffis VW, Stedinger JR (2007) Evolution of flood frequency analysis with Bulletin 17. *J Hydrol Eng* 12:283–297
- Gudmundsson L, Wagener T, Tallaksen LM, Engeland K (2012) Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe. *Water Resour Res*. <https://doi.org/10.1029/2011WR010911>
- Gutjahr O, Heinemann G (2013) Comparing precipitation bias correction methods for high-resolution regional climate simulations using COSMO-CLM: effects on extreme values and climate change signal. *Theoret Appl Climatol* 114(3–4):511–529. <https://doi.org/10.1007/s00704-013-0834-z>
- Hamon WR (1964) Computation of direct runoff amounts from storm rainfall. In: General assembly of Berkeley, symposium on surface waters: Vol. Extract of, pp 52–62
- Hattermann FF, Vetter T, Breuer L, Su B, Daggupati P, Donnelly C, Fekete B, Florke F, Gosling SN, Hoffmann P, Liersch S, Masaki Y, Motovilov Y, Muller C, Samaniego L, Stacke T, Wada Y, Yang T, Krysanova V (2018) Sources of uncertainty in hydrological climate impact assessment: a cross-scale study. *Environ Res Lett*. <https://doi.org/10.1088/1748-9326/aa9938>
- Her Y, Yoo SH, Cho J, Hwang S, Jeong J, Seong C (2019) Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions. *Sci Rep* 9(1):1–22. <https://doi.org/10.1038/s41598-019-41334-7>
- Intergovernmental Panel On Climate Change—IPCC (2013) Summary for policymakers. In: Climate change 2013. The physical science basis. Cambridge University Press, Cambridge, p 571
- Joseph J, Ghosh S, Pathak A, Sahai AK (2018) Hydrologic impacts of climate change: comparisons between hydrological parameter uncertainty and climate model uncertainty. *J Hydrol* 566:1–22. <https://doi.org/10.1016/j.jhydrol.2018.08.080>
- Kay AL, Davies HN, Bell VA, Jones RG (2009) Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Clim Change* 92(1–2):41–63. <https://doi.org/10.1007/s10584-008-9471-4>
- Keast D, Ellison J (2013) Magnitude frequency analysis of small floods using the annual and partial series. *Water (Switzerland)* 5(4):1816–1829. <https://doi.org/10.3390/w5041816>
- Keith JB (2001) Rainfall-runoff modelling: the primer. John Wiley & Sons Ltd, Chichester
- Kingston DG, Thompson JR, Kite G (2011) Uncertainty in climate change projections of discharge for the Mekong River Basin. *Hydrol Earth Syst Sci* 15(5):1459–1471. <https://doi.org/10.5194/hess-15-1459-2011>
- Knutti R, Sedláček J (2013) Robustness and uncertainties in the new CMIP5 climate model projections. *Nat Clim Change* 3(4):369–373. <https://doi.org/10.1038/nclimate1716>
- Krinner G, Flanner MG (2018) Striking stationarity of large-scale climate model bias patterns under strong climate change. *Proc Natl Acad Sci USA* 115(38):9462–9466. <https://doi.org/10.1073/pnas.1807912115>
- Lafon T, Dadson S, Buys G, Prudhomme C (2013) Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *Int J Climatol* 33(6):1367–1381. <https://doi.org/10.1002/joc.3518>

- Lawrence D (2020) Uncertainty introduced by flood frequency analysis in projections for changes in flood magnitudes under a future climate in Norway. *J Hydrol Reg Stud* 28:100675. <https://doi.org/10.1016/j.ejrh.2020.100675>
- Madison AM, Ruark MD, Stuntebeck TD, Komiskey MJ, Good LW, Drummy N, Cooley ET (2014) Characterizing phosphorus dynamics in tile-drained agricultural fields of eastern Wisconsin. *J Hydrol* 519(Part A(0)):892–901. <https://doi.org/10.1016/j.jhydrol.2014.08.016>
- Mendez M, Maathuis B, Hein-Griggs D, Alvarado-Gamboa LF (2020) Performance evaluation of bias correction methods for climate change monthly precipitation projections over Costa Rica. *Water (Switzerland)* 12(2):482. <https://doi.org/10.3390/w12020482>
- Meresa HK (2020) River flow characteristics and changes under the influence of varying climate conditions. *Nat Res Model*. <https://doi.org/10.1111/nrm.12242>
- Meresa HK, Gatachew MT (2019) Climate change impact on river flow extremes in the Upper Blue Nile River basin. *J Water Clim Change* 10(4):759–781. <https://doi.org/10.2166/wcc.2018.154>
- Meresa HK, Romanowicz RJ (2017) The critical role of uncertainty in projections of hydrological extremes. *Hydrol Earth Syst Sci*. <https://doi.org/10.5194/hess-21-4245-2017>
- Meresa HK, Romanowicz RJ, Napiorkowski JJ (2017) Understanding changes and trends in projected hydroclimatic indices in selected Norwegian and Polish catchments. *Acta Geophys* 65(4):829–848. <https://doi.org/10.1007/s11600-017-0062-5>
- Meresa H, Tischbein B, Mendela J, Demoz R, Abreha T, Weldemichael M, Ogbu K (2021) The role of input and hydrological parameters uncertainties in extreme hydrological simulations. *Nat Resour Model*. <https://doi.org/10.1111/nrm.12320>
- Mockler E, Bruen M, Desta M, Misstear B (2014) Pathways project final report volume 4: catchment modelling tool (STRIVE Report). <https://doi.org/10.13140/RG.2.1.2265.0723>
- Mockler EM, O'Loughlin FE, Bruen M (2016) Understanding hydrological flow paths in conceptual catchment models using uncertainty and sensitivity analysis. *Comput Geosci* 90:66–77. <https://doi.org/10.1016/j.cageo.2015.08.015>
- Nielsen SA, Hansen E (1973) Numerical simulation of the rainfall-runoff process on a daily basis. *Nord Hydrol* 4:171–190. <https://doi.org/10.2166/nh.1973.013>
- O'Brien RJ, Misstear BD, Gill LW, Deakin JL, Flynn R (2013) Developing an integrated hydrograph separation and lumped modelling approach to quantifying hydrological pathways in Irish river catchments. *J Hydrol* 486:259–270. <https://doi.org/10.1016/j.jhydrol.2013.01.034>
- Osuch M, Lawrence D, Meresa HK, Napiorkowski JJ, Romanowicz RJ (2017) Projected changes in flood indices in selected catchments in Poland in the 21st century. *Stoch Environ Res Risk Assess* 31(9):2435–2457. <https://doi.org/10.1007/s00477-016-1296-5>
- Pechlivanidis IG, Arheimer B, Donnelly C, Hundecha Y, Huang S, Aich V, Samaniego L, Eisner S, Shi P (2017) Analysis of hydrological extremes at different hydro-climatic regimes under present and future conditions. *Clim Change* 141(3):467–481. <https://doi.org/10.1007/s10584-016-1723-0>
- Perrin VC, Lerat J, Vaze J, Bourqui M, Hendrickx F (2012) Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. *Water Resour Res* 48(1):17. <https://doi.org/10.1029/2011WR011721>
- Piani C, Haerter JO, Coppola E (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoret Appl Climatol* 99(1–2):187–192. <https://doi.org/10.1007/s00704-009-0134-9>
- Pierce DW, Cayan DR, Maurer EP, Abatzoglou JT, Hegewisch KC (2015) Improved bias correction techniques for hydrological simulations of climate change. *J Hydrometeorol* 16(6):2421–2442. <https://doi.org/10.1175/JHM-D-14-0236.1>
- Rao AR, Delleur JW, Sarama PBS (1972) Conceptual hydrologic models for urbanizing basins. *J Hydraul Div ASCE* 98(HY7):1205–1220
- Refsgaard JC, Drews M, Jeppesen E, Madsen H, Markandya A, Olesen JE, Porter JR, Christensen JH (2013) The role of uncertainty in climate change adaptation strategies: a Danish water management example. *Mitig Adapt Strateg Glob Change* 18:337–359. <https://doi.org/10.1007/s11027-012-9366-6>
- Rojas R, Feyen L, Watkiss P (2013) Climate change and river floods in the European Union: socio-economic consequences and the costs and benefits of adaptation. *Glob Environ Change* 23(6):1737–1751. <https://doi.org/10.1016/j.gloenvcha.2013.08.006>
- Roth V, Lemann T, Zeleke G, Teklay A (2018) Effects of climate change on water resources in the upper Blue Nile Basin of Ethiopia. *Heliyon*. <https://doi.org/10.1016/j.heliyon.2018.e00771>
- Saini R, Wang G, Yu M, Kim J (2015) Phytoplankton light absorption and the package effect in relation to photosynthetic and photoprotective pigments in the northern tip of Antarctic Peninsula. *J Geophys Res*. <https://doi.org/10.1002/2014JD022599>. Received

- Serdeczny O, Adams S, Coumou D, Hare W, Perrette M (2016) Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions. *Reg Environ Change*. <https://doi.org/10.1007/s10113-015-0910-2>
- Shawul AA, Chakma S, Melesse AM (2019) The response of water balance components to land cover change based on hydrologic modeling and partial least squares regression (PLSR) analysis in the Upper Awash basin. *J Hydrol Reg Stud* 26:100640. <https://doi.org/10.1016/j.ejrh.2019.100640>
- Soriano E, Mediero L, Garjón C (2019) Selection of bias correction methods to assess the impact of climate change on flood frequency curves. *Water (Switzerland)* 11(11):2266. <https://doi.org/10.3390/w11112266>
- Sun WC, Ishidaira H, Bastola S (2010) Towards improving river discharge estimation in ungauged basins: calibration of rainfall-runoff models based on satellite observations of river flow width at basin outlet. *Hydrol Earth Syst Sci* 14:2011–2022
- Taye MT, Willems P, Block P (2015) Implications of climate change on hydrological extremes in the Blue Nile basin: a review. *J Hydrol Reg Stud* 4:280–293. <https://doi.org/10.1016/j.ejrh.2015.07.001>
- Tegegne G, Melesse AM, Asfaw DH, Worqlul AW (2020) Flood frequency analyses over different basin scales in the Blue Nile River basin, Ethiopia Getachew. *Hydrology* 7:44. <https://doi.org/10.3390/hydrology7030044>
- Teng J, Vaze J, Chiew FHS, Wang B, Perraud JM (2012) Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff. *J Hydrometeorol* 13(1):122–139. <https://doi.org/10.1175/JHM-D-11-058.1>
- Teutschbein C, Seibert J (2013) Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions. *Hydrol Earth Syst Sci* 17(12):5061–5077. <https://doi.org/10.5194/hess-17-5061-2013>
- Thober S, Kumar R, Wanders N, Marx A, Pan M, Rakovec O, Samaniego L, Sheffield J, Wood EF, Zink M (2018) Multi-model ensemble projections of European river floods and high flows at 1.5, 2, and 3 degrees global warming. *Environ Res Lett* 13(1):014003. <https://doi.org/10.1088/1748-9326/aa9e35>
- Woldemeskel FM, Sharma A, Sivakumar B, Mehrotra R (2014) A framework to quantify GCM uncertainties for use in impact assessment studies. *J Hydrol* 519(PB):1453–1465. <https://doi.org/10.1016/j.jhydrol.2014.09.025>
- Yan D, Werners SE, Ludwig F, Huang HQ (2015) Hydrological response to climate change: the Pearl River, China under different RCP scenarios. *J Hydrol Reg Stud* 4(PB):228–245. <https://doi.org/10.1016/j.ejrh.2015.06.006>
- Yang W, Andréasson J, Graham LP, Olsson J, Rosberg J, Wetterhall F (2010) Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrol Res* 41(3–4):211–229. <https://doi.org/10.2166/nh.2010.004>
- Zhang Q, Tang Q, Knowles JF, Livneh B (2019) Contribution of model parameter uncertainty to future hydrological projections. *Hydrol Earth Syst Sci Discuss*. <https://doi.org/10.5194/hess-2019-52>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Hadush Meresa^{1,2} · Bernhard Tischbein¹ · Tewodros Mekonnen²

✉ Hadush Meresa
kidane.hadush@gmail.com

¹ Ecology and Natural Resources, Center for Development Research, University of Bonn, Bonn, Germany

² Ethiopian Construction Design and Supervision Work Corporation, Addis Ababa, Ethiopia