



OPEN

DATA DESCRIPTOR

# Bias-corrected climate projections for South Asia from Coupled Model Intercomparison Project-6

Vimal Mishra <sup>1,2</sup>✉, Udit Bhatia <sup>1</sup> & Amar Deep Tiwari<sup>1</sup>

Climate change is likely to pose enormous challenges for agriculture, water resources, infrastructure, and livelihood of millions of people living in South Asia. Here, we develop daily bias-corrected data of precipitation, maximum and minimum temperatures at 0.25° spatial resolution for South Asia (India, Pakistan, Bangladesh, Nepal, Bhutan, and Sri Lanka) and 18 river basins located in the Indian sub-continent. The bias-corrected dataset is developed using Empirical Quantile Mapping (EQM) for the historic (1951–2014) and projected (2015–2100) climate for the four scenarios (SSP126, SSP245, SSP370, SSP585) using output from 13 General Circulation Models (GCMs) from Coupled Model Intercomparison Project-6 (CMIP6). The bias-corrected dataset was evaluated against the observations for both mean and extremes of precipitation, maximum and minimum temperatures. Bias corrected projections from 13 CMIP6-GCMs project a warmer (3–5°C) and wetter (13–30%) climate in South Asia in the 21<sup>st</sup> century. The bias-corrected projections from CMIP6-GCMs can be used for climate change impact assessment in South Asia and hydrologic impact assessment in the sub-continental river basins.

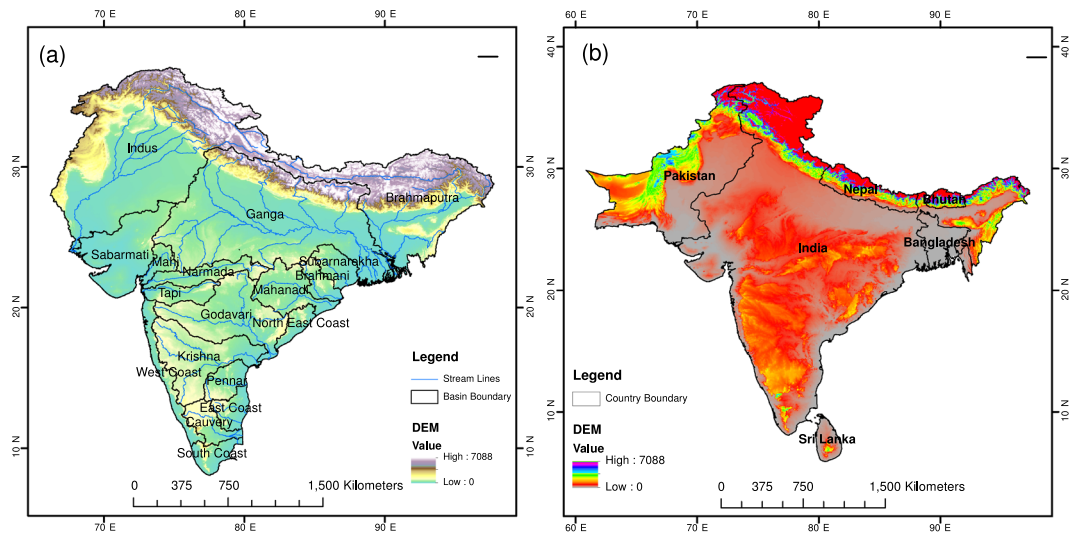
## Background & Summary

South Asia is one of the most densely populated regions of the world. A majority of the population in South Asia depends on agriculture for their livelihood. South Asia is among the global hot spots that are likely to face the detrimental impacts of climate change<sup>1,2</sup>. Considerable changes in precipitation and temperature are projected in South Asia that will have implications for water resources and agriculture<sup>3–6</sup>. The risk of floods and droughts are likely to increase in South Asia under the warming climate<sup>7–12</sup>. Both recent droughts and floods have affected a large population and caused enormous damage to agriculture and infrastructure in South Asia<sup>13–16</sup>. Similarly, the frequency and intensity of severe heatwaves have increased in South Asia and projected to increase in the future<sup>17–21</sup>. Overall, the frequency of both precipitation and temperature extremes has considerably increased in the past decades and likely to rise further under the warming climate<sup>18,22</sup>.

Projections from the General Circulation Models (GCMs) play a vital role in understanding the future changes in climate. However, spatial resolution at which GCMs are run is often too coarse to get reliable projections at the regional and local scale<sup>23</sup>. Precipitation and temperature projections at higher spatial resolution are required for the climate impact assessments<sup>24–26</sup>. Moreover, precipitation and temperature from the GCMs have a bias due to their coarse resolution or model parameterizations<sup>27,28</sup>. Therefore, for the assessment of the climate change and its impacts on different sectors (e.g., water resources, agriculture), bias-correction is required<sup>23,29–34</sup>. Both statistical and dynamical approaches are used for downscaling and bias correction of climate change projections from GCMs. Statistical approaches are based on the distribution and relationship between the observed and projected data for the historical period<sup>33,34</sup>. On the other hand, dynamical downscaling approaches are based on regional climate model forced with the boundary conditions from the coarse resolution GCMs<sup>35,36</sup>. Both statistical and dynamical downscaling approaches have limitations<sup>37,38</sup>. The primary limitation of the dynamical downscaling is related to the requirement of computational efforts to run the regional climate models at higher spatial and temporal resolution<sup>27,39</sup>. Moreover, dynamical downscaling may not remove the bias in climate variables, which might require corrections based on the statistical approaches<sup>39</sup>. Given these limitations, statistical bias correction approaches are widely used in climate change impact assessments<sup>40,41</sup>.

Considering the climate change impacts in South Asia, we develop a bias-corrected dataset of daily precipitation, maximum and minimum temperatures using output from 13 GCMs that participated in the Coupled Model

<sup>1</sup>Civil Engineering, Indian Institute of Technology (IIT) Gandhinagar, Palaj, Gujarat, 382355, India. <sup>2</sup>Earth Sciences, Indian Institute of Technology (IIT) Gandhinagar, Palaj, Gujarat, 382355, India. ✉e-mail: [vmishra@iitgn.ac.in](mailto:vmishra@iitgn.ac.in)



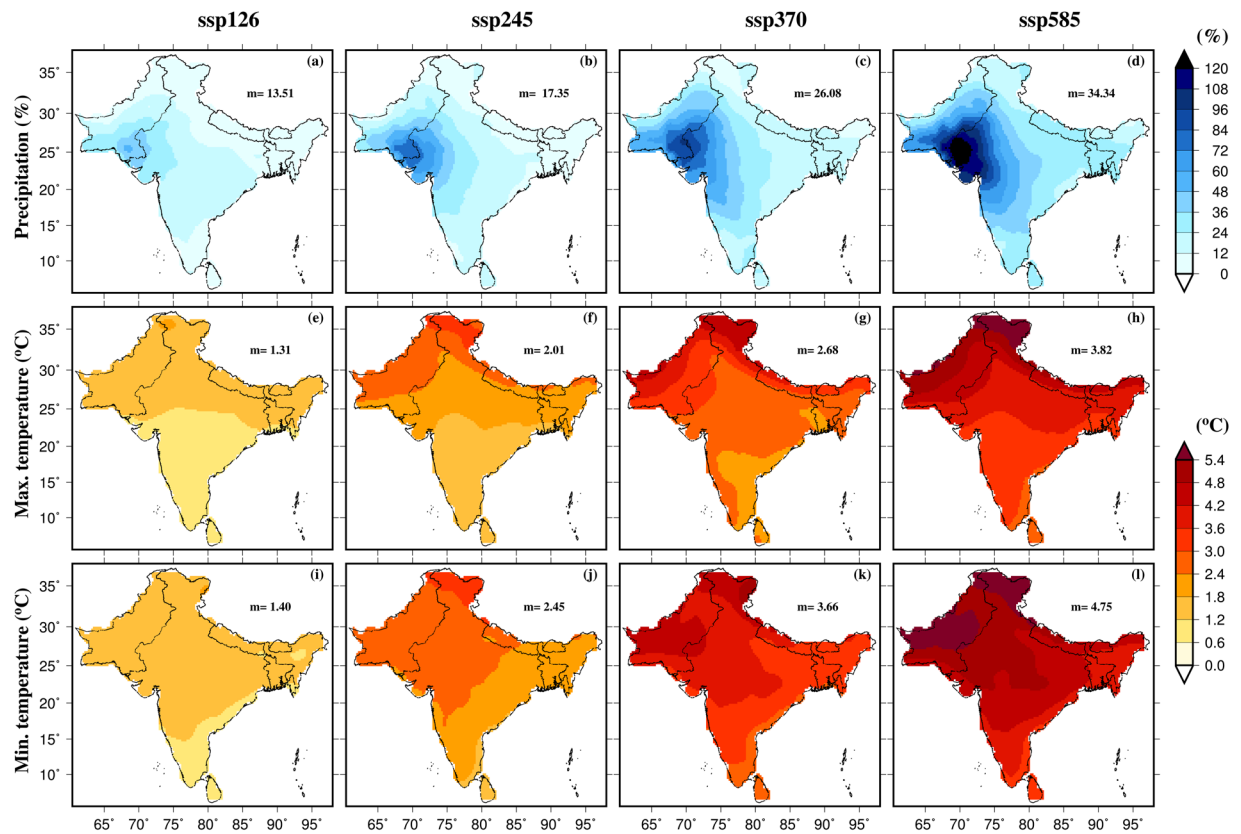
**Fig. 1** Geographical domains for bias-corrected CMIP6 projections. **(a)** Indian Subcontinent river basin boundaries (black) with the streamlines (blue). Topography in the color scale is shown in the background. Names of the sub-continental river basins are written within the basin boundaries. **(b)** South Asian country boundaries (black) where topography in the color scale is shown in the background.

Intercomparison Project-6 (CMIP6). The 13 GCMs were selected based on the availability of daily precipitation, maximum and minimum temperatures for the historical and four scenarios (SSP126, SSP245, SSP370, SSP585). We used empirical quantile mapping (EQM) to develop bias-corrected data at daily temporal and  $0.25^\circ$  spatial resolution for six countries in South Asia (India, Pakistan, Bangladesh, Nepal, Bhutan, and Sri Lanka). Also, the bias-corrected projections are developed for 18 sub-continental river basins. The bias-corrected projections from 13 CMIP-GCMs can be used for estimating the projected changes in mean and extreme climate in South Asia. Bias corrected data for 18 sub-continental river basins can be used to develop hydrologic projections using hydrological models.

## Methods

Bias-corrected projections were developed for South Asia (India, Pakistan, Bangladesh, Nepal, Bhutan, and Sri Lanka) and the 18 Indian sub-continental river basins (Fig. 1). We used basin boundaries in the Indian sub-continent from Shah and Mishra<sup>42</sup> [Fig. 1]. We obtained observed daily gridded precipitation, minimum and maximum temperatures for South Asia for the 1951–2018 period. Daily precipitation at  $0.25^\circ$  was obtained from the India Meteorological Department (IMD) for the Indian region<sup>43</sup>. Pai *et al.*<sup>43</sup> developed gridded daily precipitation for India using station observations from more than 6000 stations located across India. The precipitation captures critical features of the Indian summer monsoon, including higher rainfall in the Western Ghats and northeastern India and lower rainfall in the semi-arid and arid regions of western India. Besides, gridded precipitation captures the orographic rain in the Western Ghats and foothills of Himalaya. The gridded precipitation data from IMD has been used for various hydroclimatic applications<sup>43,44,45</sup>. Gridded daily maximum and minimum temperatures from IMD were developed using station-based observations from more than 350 stations located across India<sup>46</sup>. There is bias in temperature observations from IMD in the Himalayan region, which can be attributed to sparse station density<sup>44,47</sup>. Gridded precipitation and maximum and minimum temperatures were obtained from Sheffield *et al.*<sup>48</sup> for the regions outside India. Datasets from Sheffield *et al.*<sup>48</sup> are available at  $0.25^\circ$  spatial and daily temporal resolutions. Consistency between IMD and Sheffield *et al.*<sup>48</sup> dataset was checked in Shah and Mishra<sup>42</sup>, who reported that Sheffield *et al.*<sup>48</sup> dataset has a good agreement with the IMD observations. Nonetheless, we used IMD gridded dataset for the Indian region and Sheffield *et al.*<sup>48</sup> for outside India for bias correction of projections from CMIP6 as the IMD data is widely used for hydroclimatic studies in India. We used gridded observations for bias correction as station data are not available.

We obtained daily precipitation, maximum and minimum temperatures from 13 CMIP6-GCMs from <https://esgf-node.llnl.gov/search/cmip6/>. All the three variables and for all the scenarios were available only for these 13 GCMs. Therefore, we restricted bias-correction to only these models. Precipitation, maximum and minimum temperatures from CMIP6-GCMs are available at different spatial resolutions (Table S1). For instance, the spatial resolution of the CMIP6 projection varies from  $0.7^\circ$  (EC-Earth3) to more than  $2^\circ$  (CanESM5). All the three variables were selected for the historical (1850–2014), ssp126 (2015–2100), ssp245 (2015–2100), ssp370 (2015–2100), and ssp585 (2015–2100) scenarios under r1i1p1f1 initial condition at daily time scale<sup>49</sup>. The scenarios used in the CMIP6 combine Shared Socioeconomic Pathways (SSP) and target radiative forcing levels at the end of the 21<sup>st</sup> century<sup>50</sup>. For instance, SSP126 indicates SSP-1 and target radiative forcing at the end of the 21<sup>st</sup> century 2.6 Watt/m<sup>2</sup>. Therefore, SSP126 is a mitigation scenario. On the other hand, SSP585 is based on the emission scenario considering SSP-5 and radiative forcing of 8.5 Watt/m<sup>2</sup> at the end of the 21<sup>st</sup> century<sup>50</sup>. Further details on the scenarios used in the CMIP6 can be obtained from Gidden *et al.*<sup>50</sup>. We regridded all the variables from CMIP6 to



**Fig. 2** Projections of precipitation, maximum and minimum temperatures for the end of 21<sup>st</sup> century using raw output from CMIP6-GCMs. (a–d) Multimodel ensemble mean projected change in mean annual precipitation (%) for the Far (2074–2100) with respect to the historical period (1988–2014), (e–h) same as (a–d) but for the mean annual maximum temperature, (i–l) same as (a–d) but for the mean annual minimum temperature. Median of the multimodel ensemble mean precipitation; maximum and minimum temperatures are shown in each panel. Projected changes were estimated for the four scenarios (SSP126, SSP245, SSP370, and SSP585) against the historical period.

1° spatial resolution to make them consistent. However, the effect of regridding using bilinear interpolation was checked by comparing the gridded datasets against the raw data for all-India mean of precipitation, maximum and minimum temperatures. We did not find any considerable differences in the all-India averaged precipitation and temperature using regridded and raw output from the GCMs (Fig. S1).

Outputs of the various atmospheric (e.g., maximum and minimum temperatures, and precipitation) variables obtained from GCMs are known to exhibit systematic biases (Fig. S2). Hence, these outputs need to be bias-corrected to produce reliable estimates at regional and local scales for climate impact assessment. To achieve this, statistical transformations that attempt to find a function that maps the model output to a new distribution such that the resulting distribution matches that of observations. In general, this transformation can be formulated as Piani *et al.*<sup>30</sup>:

$$x_m^o = f(x_m) \quad (1)$$

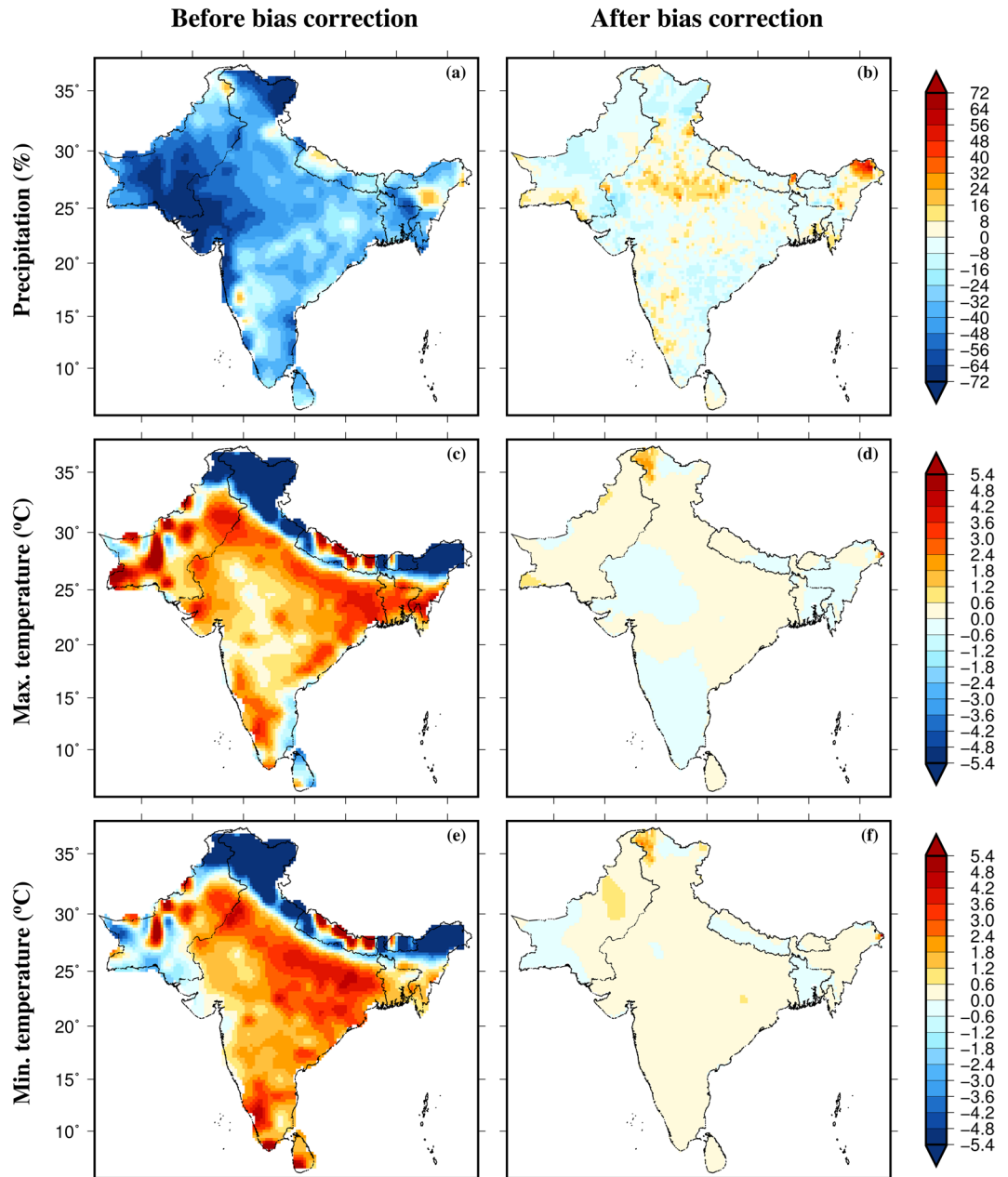
where  $x_m^o$  is the bias-corrected model output. If the statistical distribution of  $x_m$  and  $x_o$  are known, the transformation can be written as:

$$x_m^o = F_o^{-1}(F_m(x_m)) \quad (2)$$

where  $F_m$  and  $F_o$  are the Cumulative Distribution Functions (CDFs) of  $x_m$  and  $x_o$  respectively.

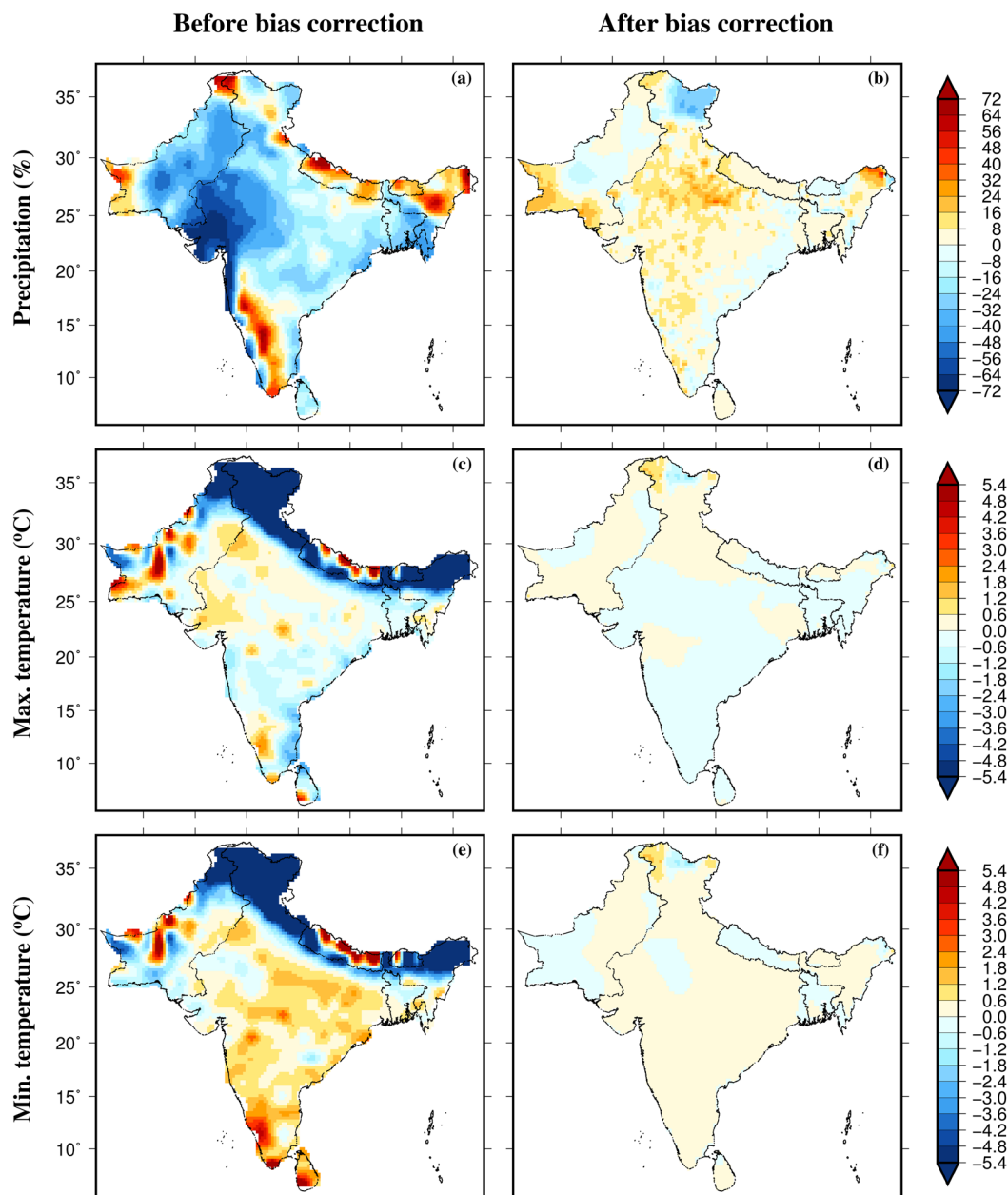
In Empirical Quantile Mapping (EQM<sup>44,51</sup>), instead of assuming parametric distributions, empirical CDFs<sup>34,52,53</sup> are estimated from the percentiles calculated from  $x_m$  and  $x_o$ . As a result, EQM and its variants can be applied to both temperature and precipitation even if their underlying distributions are different and hence recommended for statistical bias correction<sup>54</sup>.

In the context of statistical downscaling, since the observations are at a higher resolution than models, EQM on bilinearly interpolated model outputs at observation resolution is often used to address the scale mismatch and generate post-processed model outputs<sup>44</sup>. We choose non-parametric transformation approaches over the parametric approaches as has shown better skills in the comparison to parametric methods in reducing biases from GCM as well as Regional Climate Model (RCM) outputs<sup>55</sup>.



**Fig. 3** Multimodel ensemble mean bias in precipitation, maximum and minimum temperatures in 13 CMIP6-GCMs. **(a)** Bias (%) in mean annual precipitation for the historical period (1985–2014), **(b)** bias in mean annual precipitation (%) after the bias correction, **(c,d)** Bias (°C) in mean annual maximum temperature before and after bias correction, and **(e,f)** bias (°C) in mean annual minimum temperature before and after bias correction.

We used EQM to statistically downscale the daily maximum and minimum temperatures, and precipitation for South Asia and Indian sub-continental river basins (Fig. 1). We use the outputs ( $x_m$ ) from 13 CMIP6-GCMs (Table S1), which are available at different resolutions (Table S1). Observations for the three variables at the resolution of 0.25-degree are obtained from the IMD, Pai *et al.*<sup>43</sup> for Indian Region and Sheffield *et al.*<sup>48</sup> for grid-points within and outside India, respectively. We used the 1951–2014 period to obtain the transformation function to map the distribution of  $x_m$  to  $x_o$ . For precipitation, the drizzle effect is corrected by using a wet day threshold of 1 mm/day<sup>30,55</sup>. If the values from model projections are larger (smaller) than the training values used to estimate the empirical CDF, the correction found for the highest (lowest) quantile of the training period is used. We used mapped transformation to bias correct the outputs for the historical period and the SSP126, SSP245, SSP370, and SSP585 scenarios for the 2015–2100 period for all the three variables. Raw and bias-corrected data for INM-CM5 is shown against the observed maximum temperature for a randomly selected grid in the Indian subcontinent (Fig. S1). Quantile mapping based statistical bias correction has been widely used, and its performance was found to be satisfactory in comparison to the other methods<sup>25,34,56</sup>.



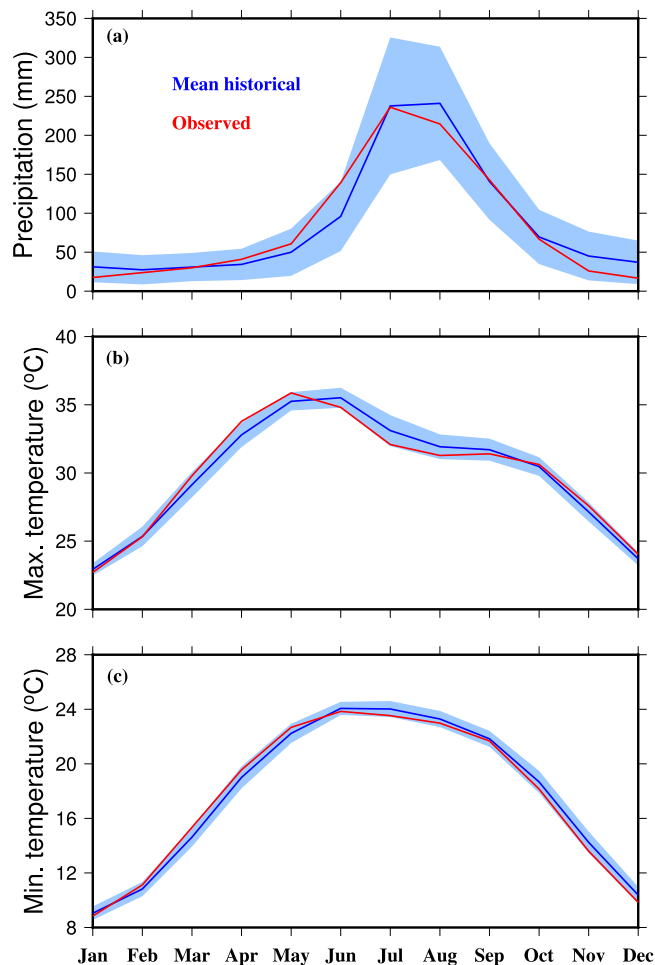
**Fig. 4** Multimodel ensemble mean bias in the 90<sup>th</sup> percentile of precipitation, maximum and minimum temperatures in 13 CMIP6-GCMs. (a) Bias (%) in extreme precipitation for the historical period (1985–2014), (b) bias in extreme precipitation (%) after the bias correction, (c,d) Bias (°C) in extreme maximum temperature before and after bias correction, and (e,f) bias (°C) in extreme minimum temperature before and after bias correction. The 90<sup>th</sup> percentile of daily precipitation was estimated using rainy days with precipitation more than 1 mm.

### Data Records

Bias corrected daily precipitation, maximum and minimum temperatures are available for the 13 GCMs (Table S1) for the historical (1951–2014) and future (2015–2100) periods. We also provide the reference gridded observed data of daily precipitation, maximum and minimum temperatures that were used for the bias correction. Projections for the future are available for the four scenarios (SSP126, SSP245, SSP370, and SSP585) for South India (India, Pakistan, Bangladesh, Sri Lanka, Bhutan, and Nepal) and Indian sub-continental river basins (Fig. 1). The basin wise dataset<sup>57</sup> and country wise dataset<sup>58</sup> have been made available through Zenodo. Details on the data format can be obtained from a readme file provided at the above link.

### Technical Validation

First, we estimated the projected changes in mean annual precipitation, maximum and minimum temperatures using the raw data from the CMIP6-GCMs (Fig. 2). The projected changes were estimated for each GCM for the late 21<sup>st</sup> century (2074–2100) against the historical reference period (1988–2014). Then, the multimodel ensemble mean of the projected changes from all the 13 CMIP6-GCMs was taken. The multimodel ensemble mean annual

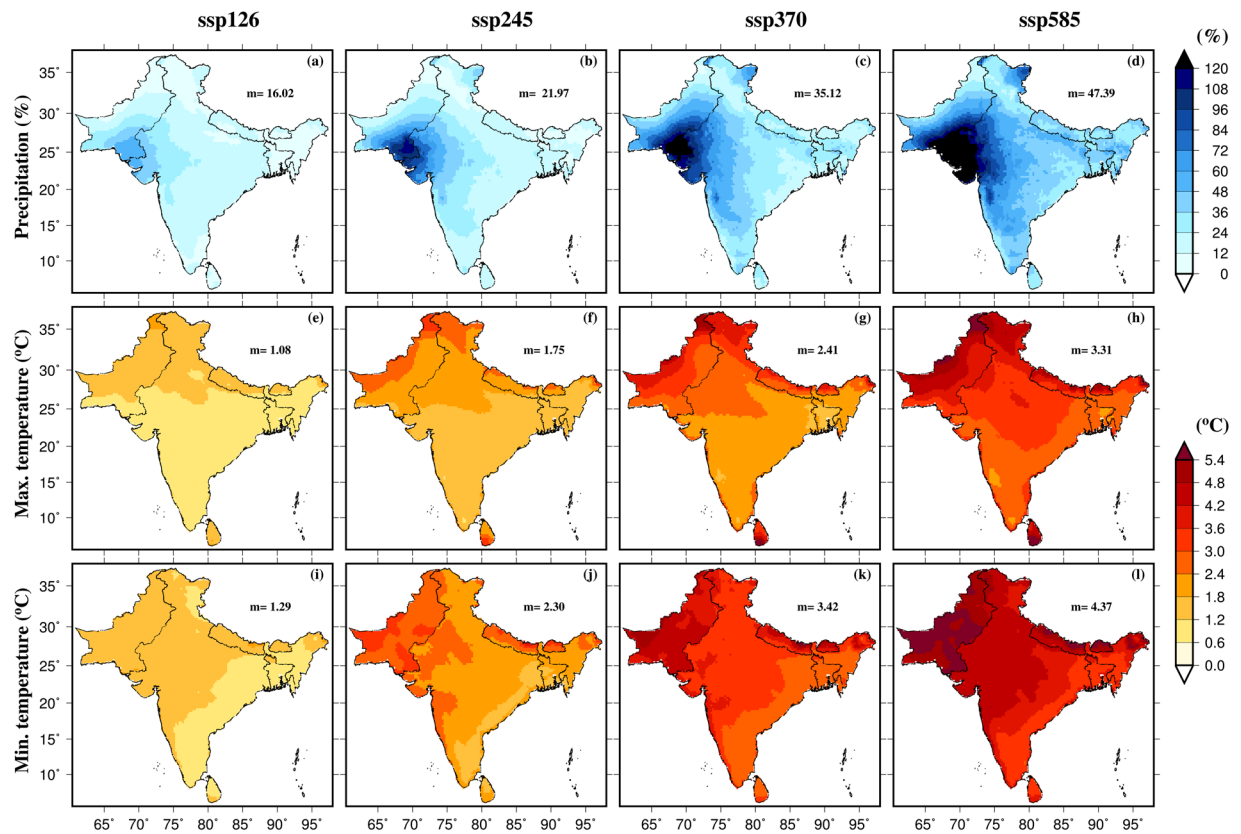


**Fig. 5** Seasonal cycle of bias-corrected precipitation, maximum and minimum temperatures. Comparison of the multimodel model ensemble (blue) mean seasonal cycle of bias-corrected (a) precipitation, (b) maximum temperature, and (c) minimum temperature against the observations for the 1985–2014 period (red). The shaded area represents uncertainty (one standard deviation) of all 13 CMIP6-GCMs.

precipitation is projected to increase in South Asia under the projected future climate (Fig. 2a–d). The projected increase in precipitation in South Asia under the future climate varies with the scenario considered. For instance, under the high-emission scenario (SSP585), a considerably higher increase (more than 30%) in the multimodel ensemble mean is projected in comparison to the low emission (SSP126) scenario (less than 13%). Similarly, there are regional differences in the projections of precipitation from the CMIP6-GCMs. For example, a more substantial increase in the multimodel ensemble mean precipitation is projected for the semi-arid and arid regions of western South Asia than the other regions (Fig. 1a–d). Similar to rainfall, mean annual maximum and minimum temperatures are projected to rise substantially in South Asia under the future climate (Fig. 2). Projected changes in mean annual minimum temperature are generally greater than the changes in mean annual maximum temperature. As expected, the high emission scenario (SSP585) will lead to a much higher rise in temperatures than the low emission scenario of SSP126 (Fig. 2).

The raw datasets of precipitation, maximum and minimum temperatures can be used to estimate the projected changes in South Asia under the future climate for different scenarios. However, climate impact studies need bias-corrected projections for decision making at regional and local scales. Since the bias-corrected dataset is consistent with observation for a climatological mean period, it is easier to infer the project changes and its implications in different sectors (e.g., water resources and agriculture) for observations. We, therefore, bias-corrected precipitation, maximum and minimum temperatures for the historical (1951–2014) and future (2015–2100) periods for all the four scenarios for South Asia and Indian sub-continental river basins. The bias-corrected dataset can be used for any region or river basin in South Asia or the Indian sub-continent (Fig. 1).

We estimated the multimodel ensemble mean bias in precipitation, maximum and minimum temperatures from the 13 CMIP6-GCMs (Fig. 3). The bias in mean annual precipitation, maximum and minimum temperatures was estimated against the observations from IMD (for the Indian domain) and Sheffield *et al.*<sup>48</sup> observations (for outside India). The CMIP6-GCMs show a dry bias (15–20%) in mean annual precipitation in the majority of South Asia (Fig. 1a). On the other hand, the multimodel ensemble mean positive bias in mean annual precipitation was found in the regions located in Nepal, Pakistan, and Peninsular India (Fig. 3a). A high cold bias in both mean annual maximum and minimum temperatures were found in the Himalayan region in the CMIP6-GCMs

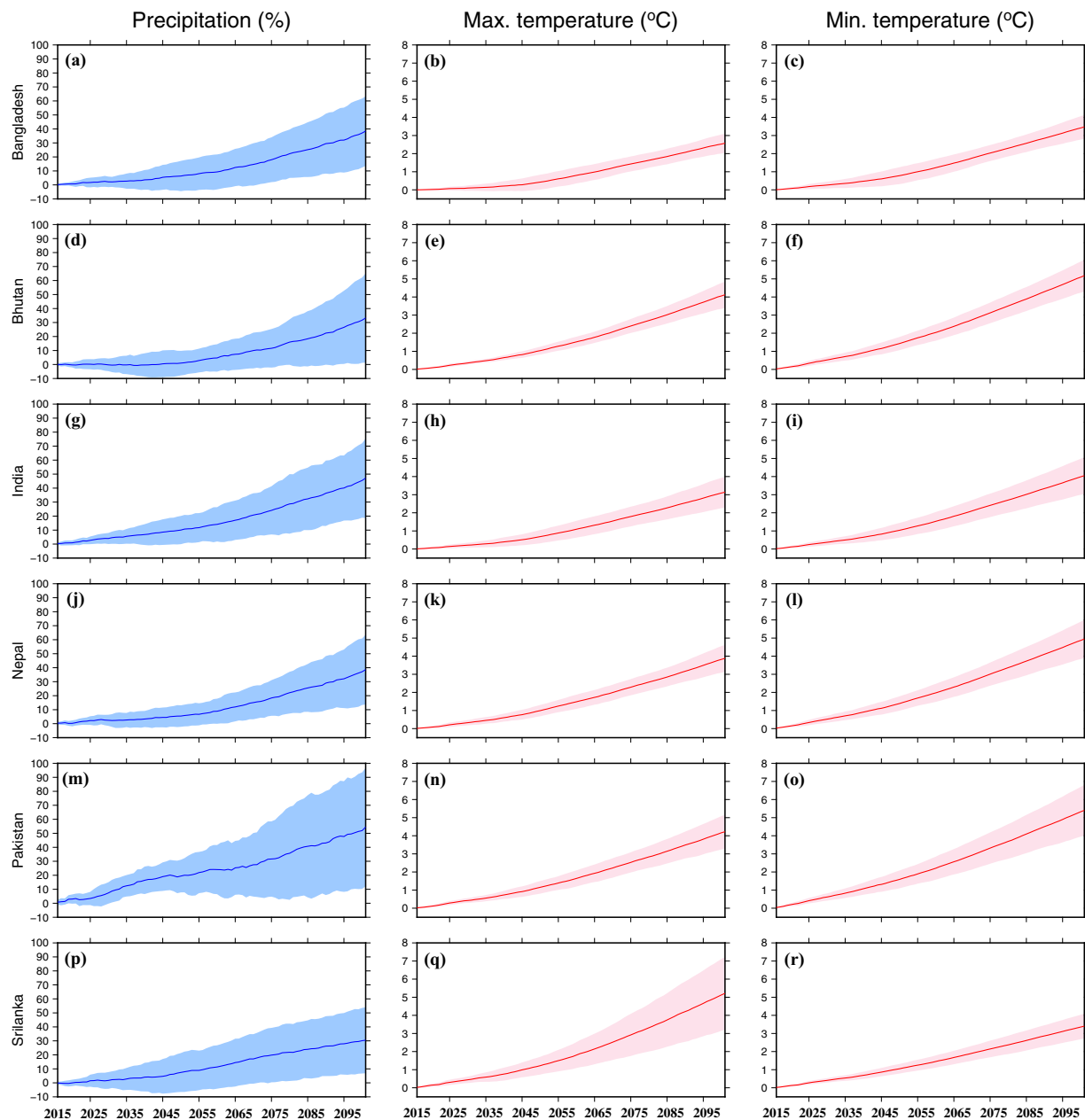


**Fig. 6** Projections of precipitation, maximum and minimum temperatures for the end of the 21<sup>st</sup> century using bias-corrected data from CMIP6-GCMs. (a–d) The multimodel ensemble mean projected change in mean annual precipitation (%) for the Far (2074–2100) with respect to the historical period (1988–2014), (e–h) same as (a–d) but for the mean annual maximum temperature, (i–l) same as (a–d) but for the mean annual minimum temperature. Median of the multimodel ensemble mean precipitation; maximum and minimum temperatures are shown in each panel. Projected changes were estimated for the four scenarios (SSP126, SSP245, SSP370, and SSP585) against the historical period.

(Fig. 3c,e). Also, CMIP6-GCMs exhibit warm bias in mean annual minimum temperature in the majority of South Asia except for the Himalayan region (Fig. 3e). We applied the EQM approach to correct the bias in the CMIP6-GCM output at daily timescale. The bias was substantially reduced after the bias correction in all the three variables for the historical (1985–2014) period (Fig. 3b–f). The reduction in bias in mean annual precipitation, maximum and minimum temperatures shows the effectiveness of our bias correction approach based on EQM.

Similar to mean annual precipitation, maximum and minimum temperatures, we estimated bias in precipitation and temperatures extremes in the raw output from the CMIP6-GCMs (Fig. 4). The 90<sup>th</sup> percentiles of precipitation of rainy days (precipitation more than 1 mm), maximum and minimum temperatures were compared for the historical period (1985–2014) from CMIP6-GCMs against the observed dataset. Consistent with mean annual precipitation, a considerable dry bias is present in extreme precipitation in CMIP6-GCMs across South Asia (Fig. 4a). We find that the CMIP6-GCMs show a warm bias in the 90<sup>th</sup> percentile of maximum and minimum temperatures across South Asia except in the Himalayan region (Fig. 4c,e). In the Himalayan region, a cool bias in CMIP6-GCMs in maximum and minimum temperature extreme was found (Fig. 4c,e). We find that the EQM based bias correction has successfully removed the bias in extreme precipitation, maximum and minimum temperatures across South Asia (Fig. 4). Therefore, the bias in both mean and extremes of precipitation, maximum and minimum temperatures were removed. Also, we compared the season cycle of bias-corrected precipitation, maximum and minimum temperatures from the CMIP6-GCMs against the observed dataset for the 1985–2014 period. Uncertainty in the bias-corrected precipitation, maximum and minimum temperatures were estimated using one standard deviation. We find that the seasonal cycle of the multimodel ensemble mean bias-corrected precipitation, maximum, and minimum temperatures compare well against the observations (Fig. 5). Moreover, the covariability of the monsoon season precipitation and air temperature is well captured by the bias-corrected dataset (Fig. S3). Overall, our results show that the EQM approach successfully corrects the bias in the CMIP6-GCMs, which can be used for climate impacts studies in South Asia. Also, the bias-corrected dataset can be used for hydrological studies in the Indian sub-continental river basins.

**Climate projections for south asia and indian sub-continental river basins.** Daily bias-corrected projections of precipitation, maximum and minimum temperatures at 0.25° from CMIP6-GCMs are developed for South Asia and the 18 Indian sub-continental river basins (Fig. 1). The projections are available for

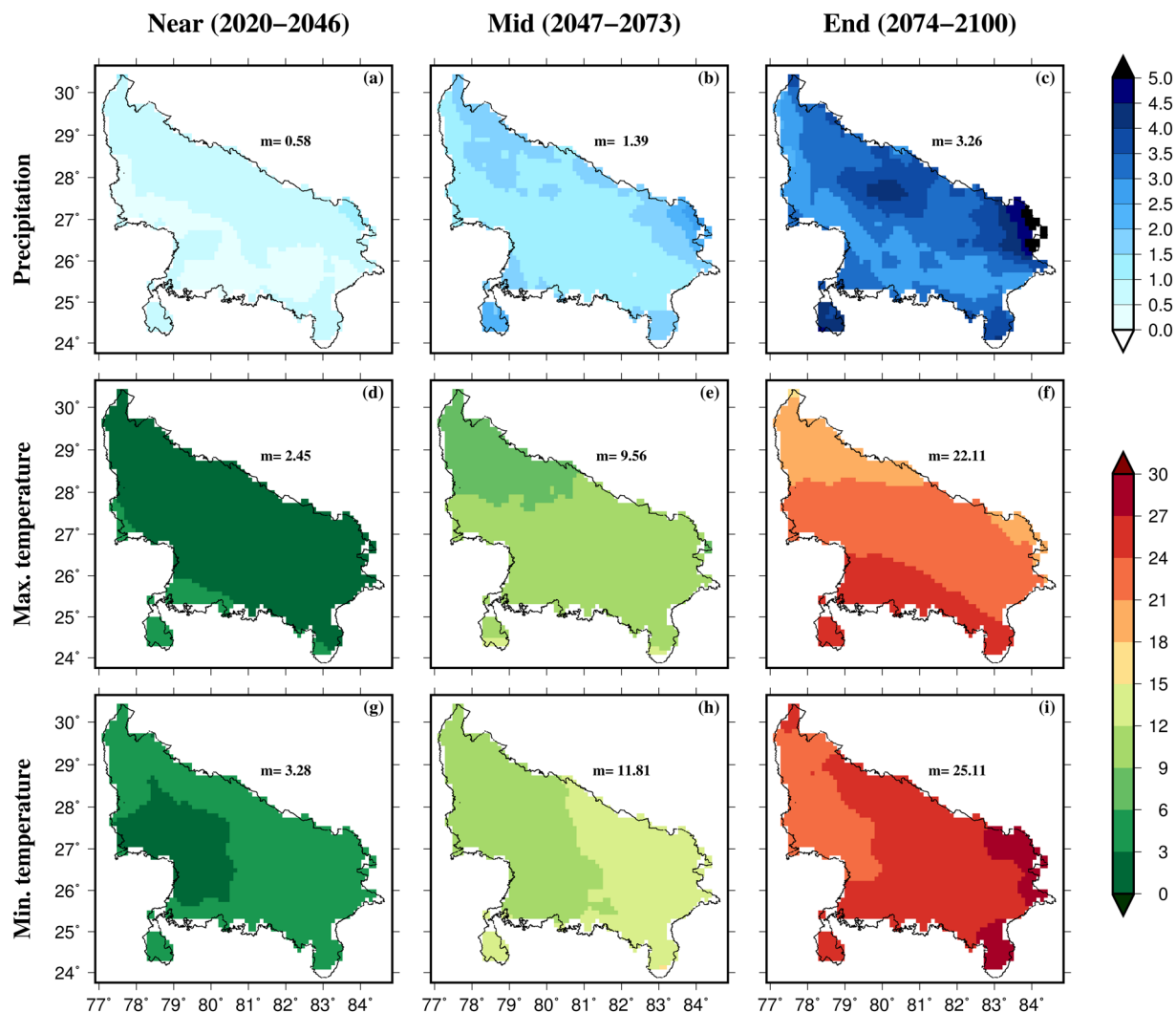


**Fig. 7** Multimodel ensemble mean change in precipitation, maximum and minimum temperatures in South Asia. Countrywise changes in the multimodel ensemble mean annual precipitation (%), maximum temperature ( $^{\circ}\text{C}$ ), and minimum temperature ( $^{\circ}\text{C}$ ) estimated using a 30-year moving window against the historical reference period of 1985–2014. The shaded region shows uncertainty (estimated using one standard deviation) based on 13 CMIP6-GCMs.

the historical (1951–2014) and future (2015–2100) periods. We estimated projected changes in precipitation, maximum and minimum temperatures in the late 21<sup>st</sup> century from the bias-corrected dataset against the historical reference (1988–2014) period for all the scenarios (SSP126, SSP245, SSP370, and SSP585) [Fig. 6]. Our bias-corrected precipitation projections show consistent spatial patterns that were observed in the raw CMIP6-GCMs (Fig. 2). For instance, a larger increase in mean annual precipitation was found in the western parts of South Asia in both raw and bias-corrected datasets (Fig. 6a–d). A considerably large increase in mean annual precipitation is projected under SSP585 (median 23%) than under SSP126 (median 12%) [Fig. 6a–d]. The ensemble mean median change in maximum temperature is projected to be 1.3  $^{\circ}\text{C}$  in SSP126 and 2.2  $^{\circ}\text{C}$  in SSP585 (Fig. 6e–h). Similarly, the ensemble mean minimum temperature is projected to rise significantly across South Asia with a median increase of more than 3  $^{\circ}\text{C}$  increase in SSP585 scenario (Fig. 6i–l).

Projected changes in precipitation, maximum and minimum temperatures were estimated using a 30-year moving window for all the six countries in South Asia under the highest emission scenario of SSP585 (Fig. 7). We considered the SSP585 scenario to estimate the projected change in precipitation, maximum and minimum

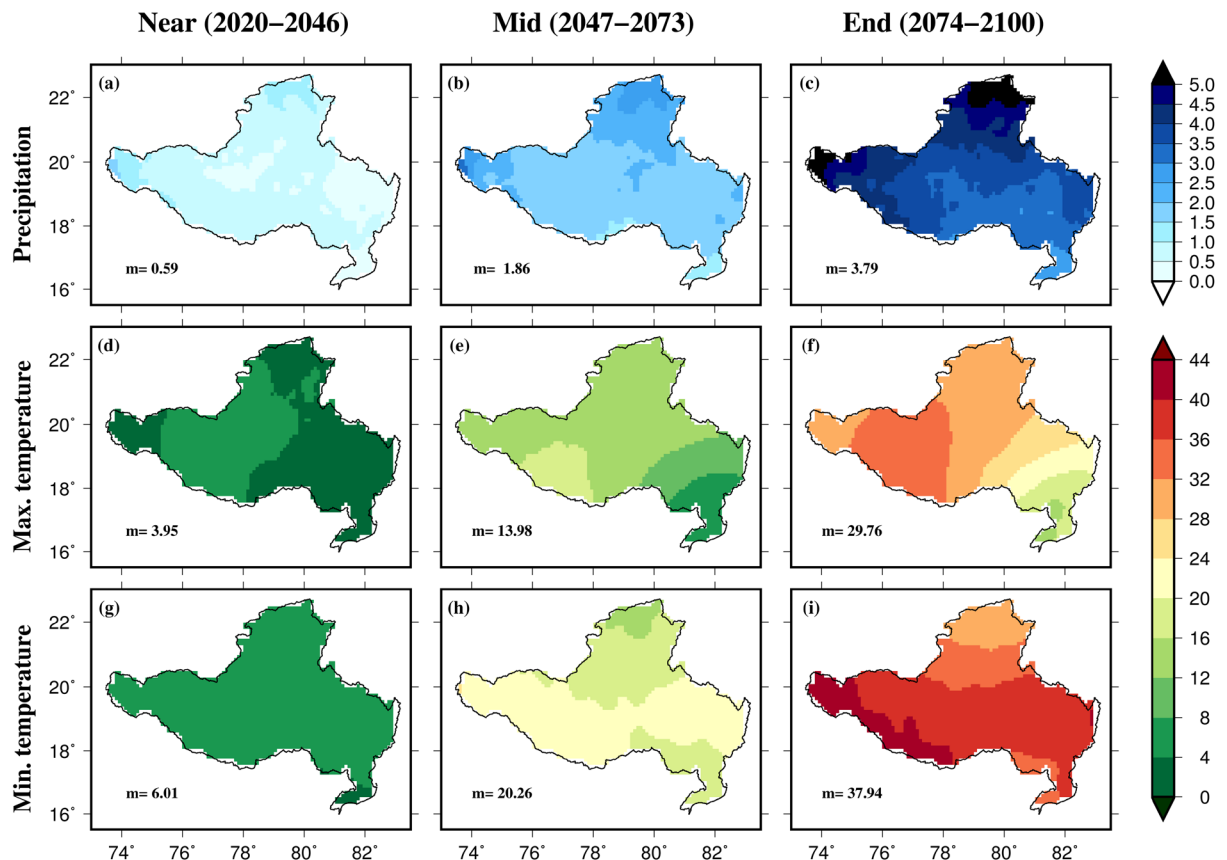




**Fig. 8** Changes in the frequency of extreme precipitation, maximum and minimum temperature in the state of Uttar Pradesh. Projected changes in the frequency of precipitation (a–c), maximum temperature (d–f), and minimum temperature (g–i) extremes estimated using 95<sup>th</sup> percentile of rainy days (precipitation more than 1 mm) and 95<sup>th</sup> percentile of maximum and minimum summer (April–May) temperatures for the state of Uttar Pradesh (India). Median frequency is shown in each panel. Changes in the frequency were estimated against the historical reference period of 1988–2014.

temperatures under the worst case (Fig. 7). Projected change in each CMIP6-GCM was estimated for each 30-year window (1986–2015, 1987–2016 ... 2071–2100) against the historical reference period of 1985–2014. Moreover, we estimated uncertainty in the bias-corrected CMIP6-GCMs using one standard deviation of projected change in the individual GCMs. The multimodel ensemble mean annual precipitation is projected to rise in all the six countries under the future climate (Fig. 7). All the six countries in South Asia are projected to experience a 20–40% rise in mean annual precipitation under the SSP585 scenario by the end of the 21<sup>st</sup> century. However, the bias-corrected precipitation projections show more uncertainty for Pakistan than the other countries (Fig. 7). Uncertainty in the bias-corrected maximum and minimum temperatures is substantially lesser than that of precipitation (Fig. 7). The multimodel ensemble mean bias-corrected mean annual maximum temperature is projected to rise by 3–4 °C by the end of the 21<sup>st</sup> century under SSP585. Moreover, the bias-corrected ensemble mean annual minimum temperature is projected to rise by 3–5 °C by the end of the 21<sup>st</sup> century (Fig. 7). We find a different level of uncertainty in mean annual precipitation, maximum and minimum temperatures for the six countries in South Asia (Fig. 7). Overall, the climate is projected to become wetter and warmer in South Asia in the future, and the magnitude of change will depend on the scenarios.

We estimated projected changes in mean annual precipitation, maximum and minimum temperatures for the six countries, and 18 sub-continental river basins for the Near (2020–2046), Mid (2047–2073), and Far (2074–2100) periods against the historical reference of 1988–2014 (Tables S2–S7). The multimodel projected changes were estimated for all the four scenarios along with the mean for the historical period (Table S2–S7). The multimodel ensemble mean bias-corrected precipitation is projected to change between 3–20% in the Near term under the SSP126 (Table S2). The most substantial increase in precipitation is projected in Pakistan, while the lowest rise



**Fig. 9** Changes in the frequency of extreme precipitation, maximum and minimum temperature in the state of Godavari basin. Projected changes in the frequency of precipitation (a–c), maximum temperature (d–f), and minimum temperature (g–i) extremes estimated using 95<sup>th</sup> percentile of rainy days (precipitation more than 1 mm) and 95<sup>th</sup> percentile of maximum and minimum summer (April–May) temperatures for the state of Uttar Pradesh (India). Median frequency is shown in each panel. Changes in the frequency were estimated against the historical reference period of 1988–2014.

is expected in Bhutan. Precipitation is projected to rise substantially in the Far term in all the countries in South Asia under the SSP585 (Table S2). The ensemble mean bias-corrected precipitation is projected to change by 31–53%, with the most considerable projected rise in Pakistan in the Far term under SSP585 (Table S2). The projected increase in mean annual maximum temperature is far lesser (0.48–0.97 °C) in the Near term under SSP126 in comparison to the Far (2.6–5.3 °C) term under the SSP585 (Table S3) in the six countries in South Asia. Our bias-corrected data based on 13 CMIP6-GCMs projected an increase of 0.7–1.3 °C) in mean annual minimum temperature in the Near term under SSP126 (Table S4).

Moreover, mean annual minimum temperature is projected to rise between 3.5 to 5.5 °C in the late 21<sup>st</sup> century under the SSP585 (Table S4). Uncertainty in the projections of bias-corrected precipitation, maximum and minimum temperatures was estimated for each country and period under all the four scenarios (Tables S2–S4). Temperature projections show lesser uncertainty than the projections of precipitation, which might have implications for hydrologic applications of the bias-corrected projections (Tables S2–S4).

We estimated projected changes in mean annual precipitation, maximum and minimum temperatures using bias-corrected data from the 13 CMIP6-GCMs for the 18 sub-continental river basins (Tables S5–S7, Fig. 1). Bias corrected projections of the three climatic variables are essential for hydrologic modelling, and the climate change impact assessment. Mean annual precipitation is projected to rise across the basins under all the scenarios in the projected future climate (Table S5). The projected rise in the mean annual precipitation is considerably higher in the SSP585 in comparison to the SSP126. The projected rise in precipitation in the sub-continental basins is higher in the Far period than the Near period.

Notwithstanding a considerable uncertainty in the precipitation projections, bias-corrected data show that precipitation is projected to rise more in the river basins located in the semi-arid/arid regions of the Indian sub-continent (Table S5). Similarly, significant warming in the mean annual maximum and minimum temperatures is projected based on the bias-corrected data from the 13 CMIP6-GCMs (Tables S6, S7). Mean annual maximum temperature is projected to rise between 2.5–4.4 °C in the Far period under SSP585 in the Indian sub-continental river basins (Table S6). Moreover, the mean annual minimum temperature is projected to rise by 3.0–5.0 °C in the Far period under SSP585 in the river basins of the Indian sub-continent (Table S7). Basin specific projections and associated uncertainty can be seen in supplemental Tables S5–S7. Overall, the bias-corrected projections can be used for the hydroclimatic impact assessment in the sub-continental river basins.

## Usage Notes

Daily bias-corrected projections of precipitation, maximum and minimum temperatures at 0.25° are essential for climate impact assessment for the administrative boundaries or at river basin scale<sup>25</sup>. We developed bias-corrected projections from 13 CMIP6-GCMs that can be used for hydroclimatic impact assessment based on mean and extremes in South Asia. Bias corrected data performs well against the mean and extremes. The dataset has been arranged based on the geographical boundaries of six countries in South Asia. Moreover, we provide separate data for each of the 18 sub-continental river basins. Daily bias-corrected projections can be used to estimate climatic indices associated to mean and extremes. For instance, daily maximum and minimum temperatures can be used to estimate projected changes under different scenarios for the crop growing seasons. Moreover, the temperature dataset can be used to estimate growing degree days (GDD)<sup>59</sup> and other indicators of extreme heat during the crop growing period<sup>60</sup>.

Similarly, daily precipitation projections can be used to estimate changes in mean and extreme precipitation for any period during the 21<sup>st</sup> century<sup>22,27</sup>. Data users can also estimate the differences in indicators and potential impacts based on the low (SSP126) and high (SSP585) emission scenarios. Most of the hydrological models require daily precipitation, maximum and minimum temperatures as the primary inputs of meteorological forcing. Therefore, hydrological models can be used with the bias-corrected projections to estimate the impacts of the projected future climate on hydrology for a river basin or a region.

As an example, we use the bias-corrected projections to estimate the frequency of precipitation and temperature extremes for an administrative region (state of Uttar Pradesh, India) and a river basin (Godavari, India) [Figs. 8, 9]. The frequency of extreme precipitation was estimated using 95<sup>th</sup> percentiles of rainy days (precipitation more than 1 mm). Similarly, the frequency of extreme hot maximum and minimum temperatures was estimated using the 95<sup>th</sup> percentile of the two hottest months (April–May) in the region. As expected, both precipitation and temperature extremes are projected to rise in Uttar Pradesh and Godavari basin under the SSP585 scenario (Figs. 8, 9). The projected rise in the frequency of precipitation and temperature extremes is higher for the Far period than the Near-term climate. Overall, daily bias-corrected CMIP6 projections can be used for multiple assessments related to climate and hydrology in one of the most populated regions of the world. As we provided the bias-corrected data for individual GCMs, users can select the GCMs that perform well in the region of interest. Moreover, the range of future projections can be estimated using the bias-corrected projections from the individual GCMs (Fig. S4). In the future, bias-corrected projections will be made available from more CMIP6-GCMs as their output becomes available. More details on the data can be found in the link and from the readme file.

## Code availability

Codes used for bias correction of CMIP6-GCMs are available through the Github link: [https://github.com/udit1408/cmip6\\_downscaling](https://github.com/udit1408/cmip6_downscaling)

Received: 11 June 2020; Accepted: 1 September 2020;

Published online: 12 October 2020

## References

1. Suarez-Gutierrez, L., Müller, W. A., Li, C. & Marotzke, J. Dynamical and thermodynamical drivers of variability in European summer heat extremes. *Clim. Dyn.* **54**, 4351–4366 (2020).
2. Diffenbaugh, N. S. & Giorgi, F. Climate change hotspots in the CMIP5 global climate model ensemble. *Clim. Change* **114**, 813–822 (2012).
3. Knox, J., Hess, T., Daccache, A. & Wheeler, T. Climate change impacts on crop productivity in Africa and South Asia. *Environ. Res. Lett.* **7**(3), 034032 (2012).
4. Turner, A. G. & Annamalai, H. Climate change and the South Asian summer monsoon. *Nature Climate Change* **2**, 587–595 (2012).
5. Lobell, D. B. & Burke, M. B. Why are agricultural impacts of climate change so uncertain? the importance of temperature relative to precipitation. *Environ. Res. Lett.* **3**(3), 034007 (2008).
6. Immerzeel, W. W., van Beek, L. P. H. & Bierkens, M. F. P. Climate change will affect the Asian water towers. *Science* **328**, 1382–5 (2010).
7. Aadhar, S. & Mishra, V. A Substantial rise in the area and population affected by dryness in South Asia under 1.5, 2.0 and 2.5C warmer worlds. *Environ. Res. Lett.* **14**, 114021 (2019).
8. Aadhar, S. & Mishra, V. Increased drought risk in South Asia under warming climate: Implications of uncertainty in potential evapotranspiration estimates. *J. Hydrometeorol.* **3** (2020).
9. Ali, H., Modi, P. & Mishra, V. Increased flood risk in Indian sub-continent under the warming climate. *Weather Clim. Extrem.* **25**, 100212 (2019).
10. Hirabayashi, Y. *et al.* Global flood risk under climate change. *Nat. Clim. Chang.* **3**, 816–821 (2013).
11. Arnell, N. W. & Gosling, S. N. The impacts of climate change on river flood risk at the global scale. *Clim. Change* **134**, 387–401 (2016).
12. Winsemius, H. C. *et al.* Global drivers of future river flood risk. *Nat. Clim. Chang.* **6**, 381–385 (2016).
13. Mishra, V., Aadhar, S., Pai, S. & Kumar, R. Supp: On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain.
14. Webster, P. J., Toma, V. E. & Kim, H. M. Were the 2010 Pakistan floods predictable? *Geophys. Res. Lett.* **38**, 1–5 (2011).
15. Hunt, K. M. R. & Menon, A. The 2018 Kerala floods: a climate change perspective. *Clim. Dyn.* **54**, 2433–2446 (2020).
16. Mishra, V. & Shah, H. L. Hydroclimatological Perspective of the Kerala Flood of 2018. *J. Geol. Soc. INDIA* **92**, 645–650 (2018).
17. Mishra, V., Mukherjee, S., Kumar, R. & Stone, D. A. Heat wave exposure in India in current, 1.5°C, and 2.0°C worlds. *Environ. Res. Lett.* **12**, 124012 (2017).
18. JA, N. *et al.* A seven-fold rise in the probability of exceeding the observed hottest summer in India in a 2°C warmer world. *Environ. Res. Lett.* **15**(4), 044028 (2020).
19. Mukherjee, S. & Mishra, V. A sixfold rise in concurrent day and night-time heatwaves in India under 2°C warming. *Sci. Rep.* **8**, 16922 (2018).
20. Im, E. S., Pal, J. S. & Eltahir, E. A. B. Deadly heat waves projected in the densely populated agricultural regions of South Asia. *Sci. Adv.* **3**, e1603322 (2017).
21. Mazdiyasi, O. *et al.* Increasing probability of mortality during Indian heat waves. *Sci. Adv.* **3**, 1–6 (2017).

22. Mukherjee, S., Aadhar, S., Stone, D. & Mishra, V. Increase in extreme precipitation events under anthropogenic warming in India. *Weather Clim. Extrem.* **20**, 45–53 (2018).
23. Christensen, J. H., Boberg, F., Christensen, O. B. & Lucas-Picher, P. On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.* **35**(20) (2008).
24. Cayan, D. R. *et al.* Future dryness in the Southwest US and the hydrology of the early 21st century drought. *Proc. Natl. Acad. Sci. USA* **107**, 21271–21276 (2010).
25. Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D. & Cayan, D. R. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol. Earth Syst. Sci.* **14**, 1125–1138 (2010).
26. Barbero, R., Fowler, H. J., Lenderink, G. & Blenkinsop, S. Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions? *Geophys. Res. Lett.* **44**, 974–983 (2017).
27. Mishra, V. *et al.* Reliability of regional and global climate models to simulate precipitation extremes over India. *J. Geophys. Res. Atmos.* **119**, 9301–9323 (2014).
28. Ashfaq, M., Rastogi, D., Mei, R., Touma, D. & Ruby Leung, L. Sources of errors in the simulation of south Asian summer monsoon in the CMIP5 GCMs. *Clim. Dyn.* **49**, 193–223 (2017).
29. Maraun, D. *et al.* Towards process-informed bias correction of climate change simulations. *Nat. Clim. Chang.* **7**, 764–773 (2017).
30. Piani, C., Haerter, J. O. & Coppola, E. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* **99**, 187–192 (2010).
31. Eisner, S., Voss, F. & Kynast, E. Statistical bias correction of global climate projections - Consequences for large scale modeling of flood flows. *Adv. Geosci.* **31**, 75–82 (2012).
32. Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Clim. Change* **62**, 189–216 (2004).
33. Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T. & Hegewisch, K. C. Improved bias correction techniques for hydrological simulations of climate change. *J. Hydrometeorol.* **16**, 2421–2442 (2015).
34. Thrasher, B., Maurer, E. P., McKellar, C. & Duffy, P. B. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrol. Earth Syst. Sci.* **16**, 3309–3314 (2012).
35. Giorgi, F. & Gutowski, W. J. Regional Dynamical Downscaling and the CORDEX Initiative. *Annu. Rev. Environ. Resour.* **40**, 467–490 (2015).
36. Mearns, L. O. *et al.* Climate change projections of the North American Regional Climate Change Assessment Program (NARCCAP). *Clim. Change* **120**, 965–975 (2013).
37. Abatzoglou, J. T. & Brown, T. J. A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.* **32**, 772–780 (2012).
38. Maurer, E. P. & Hidalgo, H. G. Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.* **14**, 1125–1138 (2008).
39. White, R. H. & Toumi, R. The limitations of bias correcting regional climate model inputs. *Geophys. Res. Lett.* **40**, 2907–2912 (2013).
40. Gutmann, E. *et al.* An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.* **50**, 7167–7186 (2014).
41. Xu, L. & Wang, A. Application of the Bias Correction and Spatial Downscaling Algorithm on the Temperature Extremes From CMIP5 Multimodel Ensembles in China. *Earth Sp. Sci.* **6**, 2508–2524 (2019).
42. Shah, H. L. & Mishra, V. Hydrologic Changes in Indian Sub-Continental River Basins (1901–2012). *J. Hydrometeorol.* **17**, 2667–2687 (2016).
43. Pai, D. S. *et al.* Development of a new high spatial resolution (0.25° × 0.25°) Long Period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam* **65**, 1–18 (2014).
44. High-resolution near real-time drought monitoring in South Asia.
45. Mishra, V. Long-term (1870–2018) drought reconstruction in context of surface water security in India. *J. Hydrol.* **580**, 124228 (2020).
46. Srivastava, A. K., Rajeevan, M. & Kshirsagar, S. R. Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region. *Atmos. Sci. Lett.* **10**, 249–254 (2009).
47. Shah, R. & Mishra, V. Evaluation of the Reanalysis Products for the Monsoon Season Droughts in India. *J. Hydrometeorol.* **15**, 1575–1591 (2014).
48. Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Clim.* **19**, 3088–3111 (2006).
49. Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
50. Gidden, M. J. *et al.* Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century. *Geosci. Model Dev.* **12**, 1443–1475 (2019).
51. Wood, A. W., Maurer, E. P., Kumar, A. & Lettenmaier, D. P. Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res. D Atmos.* **107**, 1–15 (2002).
52. Julien, B. L., T., E. H. & E., M. Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *Int. J. Climatol.* **27**, 1643–1655 (2007).
53. Jakob Themeßl, M., Gobiet, A. & Leuprecht, A. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.* **31**, 1530–1544 (2011).
54. Cannon, A. J. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Comput. Geosci.* **37**, 1277–1284 (2011).
55. Gudmundsson, L., Bremnes, J. B., Haugen, J. E. & Engen-Skaugen, T. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations &dash; A comparison of methods. *Hydrol. Earth Syst. Sci.* **16**, 3383–3390 (2012).
56. Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R. & Cannon, A. J. Downscaling extremes-an intercomparison of multiple statistical methods for present climate. *J. Clim.* **25**, 4366–4388 (2012).
57. Mishra, V., Bhatia, U. & Tiwari, A. D. Bias corrected climate projections from CMIP6 models for Indian sub-continental river basins. *Zenodo* <https://doi.org/10.5281/zenodo.3874046> (2020).
58. Mishra, V., Bhatia, U. & Tiwari, A. D. Bias Corrected Climate Projections from CMIP6 Models for South Asia. *Zenodo* <https://doi.org/10.5281/zenodo.3987736> (2020).
59. Meng, Q. & Mourshed, M. Degree-day based non-domestic building energy analytics and modelling should use building and type specific base temperatures. *Energy Build.* **155**, 260–268 (2017).
60. Ben-Ari, T. *et al.* Identifying indicators for extreme wheat and maize yield losses. *Agric. For. Meteorol.* **220**, 130–140 (2016).

## Acknowledgements

We acknowledge the data availability from India Meteorology Department (IMD) and Sheffield *et al.* (2006). Output from the CMIP6 models from <https://esgf-node.lnl.gov/projects/cmip6/> is greatly acknowledged. This work is supported by Ministry of Water Resources under National Water Mission.

### Author contributions

V.M. designed the study and wrote the manuscript. A.T. downloaded and processed CMIP6 projections. U.B. did the bias correction of CMIP6 projections.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41597-020-00681-1>.

**Correspondence** and requests for materials should be addressed to V.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files associated with this article.

© The Author(s) 2020