https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

GDDRP-CMIP6: Global Daily Runoff Projections from Bias-Corrected and Downscaled CMIP6 Forcing

Jiachao Chen[1]; Dai Yamazaki[1], Ming Pan[2]

[1] Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

[2] Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

Correspondence: Jiachao Chen (jcchen@g.ecc.u-tokyo.ac.jp)

Abstract

Global runoff projections are essential for climate-change impact assessment, yet the native runoff of CMIP6 global climate models (GCMs) is spatially coarse, biased, and generated by heterogeneous land-surface schemes across models, limiting its further hydrological applications. Here we present the Global Daily Downscaled Runoff Projections CMIP6 (GDDRP-CMIP6), a global daily runoff ensemble dataset at 0.25° covering 1950–2100, generated within a unified modelling framework by driving the calibrated Variable Infiltration Capacity (VIC) model with bias-corrected and downscaled atmospheric forcings from NEX-GDDP-CMIP6. The dataset incorporates runoff of nine GCMs under the historical experiment and four SSP scenarios, whose transient climate response covers most of the very likely range assessed in the IPCC AR6. The simulated runoff, comprising surface runoff and baseflow, is routed through the global hydrodynamic model, CaMa-Flood, and evaluated against daily observations at 2,732 GRDC stations. The derived discharge reproduces the observed seasonal cycle with a global median climatological correlation of 0.83 and shows moderate median biases in mean, high, and low flows (−5.76 %, −10.18 %, and +10.65 %). Benchmarking against the native GCM runoff and against VIC simulations driven by raw CMIP6 forcings shows that bias correction generally improves agreement with observed discharge and reduces the median inter-member spread of the historical discharge climatology by about one order of magnitude. The inter-member range of projected discharge changes is also narrowed, mainly by the common hydrological model, while the ensemble-mean projected changes are broadly consistent with those of the native GCM runoff. By providing daily total runoff from 1950 to 2100 under four SSPs within a single globally consistent framework, including many islands absent from native GCM land masks, GDDRP-CMIP6 enables directly comparable flood, drought, and water-resource assessments across regions and spatial scales. During the review process, the dataset is available at https://www.scidb.cn/en/anonymous/WkJiaUFq.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

1 Introduction

Climate change is reshaping the global hydrological cycle by shifting precipitation and evapotranspiration regimes, and the resulting changes in streamflow and their extremes are already imposing substantial costs worldwide (Chandrasekar et al., 2026; Hoang et al., 2016; Muzava et al., 2025; Singhal et al., 2025; Zam et al., 2021). Runoff is the central hydrological variable linking climatic forcing to streamflow generation, and the essential input to global river models. Its variability and trends underlie the projection of floods, droughts and shifts in water availability that affect people, ecosystems, agriculture and energy systems (Chanda et al., 2025; Sreedevi et al., 2022; Yan et al., 2019; Zhao et al., 2021). Assessing these impacts requires reliable, high-resolution runoff data (Abeysingha et al., 2025; de Alvarenga et al., 2026; Maurya et al., 2023).

The Coupled Model Intercomparison Project, now in its sixth phase (CMIP6), provides the multi-model, multi-scenario experiments on which most climate-change research depends. However, the underlying general circulation models (GCMs) carry systematic biases in precipitation and temperature (Du et al., 2022; Lun et al., 2021; Tian et al., 2023; Wang et al., 2022), which fall into three broad categories. First, mean-state errors are regionally coherent and regime-dependent: GCMs tend to underestimate precipitation in humid regions and overestimate it in arid regions (Iqbal et al., 2021; Tong et al., 2024). For example, Northern-Hemisphere summer monsoon precipitation is systematically underestimated with an ensemble-mean RMSE of about 1.66 mm d^{-1} (Cheng et al., 2024), whereas annual precipitation over arid Xinjiang is overestimated by about 89 % on average (X. Zhang et al., 2022). Second, the structure of daily precipitation is distorted: many models produce too many low-intensity wet days (Kim et al., 2020; Kurniadi et al., 2023) and underestimate extreme precipitation by 20–80 % relative to observations (Manikanta et al., 2024). Third, inter-model spread is large: projected late-century Indian summer monsoon rainfall changes vary by 1.6 ± 0.8 mm d^{-1} across CMIP6 (Sooraj et al., 2026), and the ranking of best-performing models is sensitive to the chosen metric and region (Belay et al., 2024; Roulo et al., 2026). Runoff, as a variable derived from these biased forcings, inherits their errors. Its reliability in native GCM output is further limited by structural factors, including simplified land-surface schemes and coarse grids that smooth the terrain controlling runoff generation. Raw GCM runoff is therefore unsuitable for direct use in impact assessment.

These forcing biases propagate directly into hydrological simulations. When the JULES land-surface model was driven by raw CMIP5 GCM outputs, six of nine major global basins yielded negative Nash–Sutcliffe efficiency (NSE), whereas the same model driven by reanalysis attained NSE values of 0.24–0.94, demonstrating that the failure originates in the climatic forcing rather than in the hydrological model itself (Papadimitriou et al., 2017). Uncertainty decompositions likewise identify the choice of GCM as the dominant source of uncertainty in hydrological projections, contributing roughly 60 % to 90 % of the total spread depending on region and variable (Exbrayat et al., 2014; Galavi et al., 2019; Kay et al., 2009; Lee et al., 2017; S. Zhang et al., 2022). Moreover, these climatic biases are largely retained even after dynamical downscaling (Bass et al., 2024). Bias correction is therefore a prerequisite for credible impact assessment.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

A second, related limitation of native CMIP archives is their coarse spatial resolution. CMIP5 and CMIP6 GCMs are typically discretized at 100–300 km, scales at which orographic precipitation, mesoscale convection and the sharp gradients of complex terrain are poorly resolved (Maraun et al., 2010; Prudhomme et al., 2002). Representing High Mountain Asia with smoothed, GCM-like topography reduces ridge-line precipitation by about 28 % while spuriously increasing the area-averaged amount by about 9 %, opposite-signed errors that smoothing alone cannot remove (Cannon et al., 2017). Raising resolution within the existing CMIP range only partially alleviates such problems, because model formulation matters more than resolution alone (Bjarke et al., 2024; Giorgi and Raffaele, 2022). Resolution and bias are therefore typically addressed jointly, through downscaling combined with bias correction (Wood et al., 2004). Two international efforts have shaped the available downscaled archives. The Coordinated Regional Climate Downscaling Experiment (CORDEX; Giorgi & Gutowski, 2015) provides high-resolution regional simulations, typically at 0.11°–0.44°, but its coverage is organized around individual continental domains and does not constitute a globally consistent product. The Inter-Sectoral Impact Model Intercomparison Project (ISIMIP; Lange, 2019) is global and provides bias-corrected forcing together with cross-sector impact output, but at a uniform 0.5° resolution. Currently available global datasets for assessing climate change impacts still fail to simultaneously satisfy the requirements of consistency, spatial resolution, and bias-correction, thereby constraining cross-regional and cross-scale impact assessment studies.

The NASA Earth Exchange Global Daily Downscaled Projections dataset (NEX-GDDP-CMIP6; Thrasher et al., 2022) fills this gap for the atmospheric variables. It is a bias-corrected, statistically downscaled daily archive at 0.25° covering 35 CMIP6 GCMs under multiple Shared Socioeconomic Pathways, and regional evaluations across several continents confirm that it substantially reduces the systematic biases of raw CMIP6 output (Baogang et al., 2024; Dioha et al., 2024; He et al., 2026; Jain et al., 2019; Sahany et al., 2019). However, downscaled variables do not include runoff in general, because runoff is generated from nonlinear interactions among soil moisture, snow and vegetation, incorporating strong variability, which cannot be directly disaggregated spatially. The NEX-GDDP-CMIP6 has been frequently used for impact assessments, but the hydrological applications are conducted separately for individual basins, each with its own model and calibration (Musie et al., 2020; Su et al., 2025; Usman et al., 2022). These basin-specific results fail to provide basis for cross-regional comparison. A globally consistent high-resolution runoff product driven by bias-corrected CMIP6 forcings is needed.

This study addresses this gap. We drive the calibrated Variable Infiltration Capacity (VIC) model with nine NEX-GDDP-CMIP6 ensemble members over the global land surface, generating a multi-GCM, multi-scenario daily runoff dataset at 0.25°, named Global Daily Downscaled Runoff Projections CMIP6 (GDDRP-CMIP6). The nine GCMs cover most of the assessed range of transient climate response (TCR) strengths. The simulated runoff is routed through the Catchment-based Macro-scale Floodplain (CaMa-Flood) global hydrodynamic model to obtain river discharge, which is evaluated against observations for the historical period and benchmarked against the driving GCMs to examine whether the ensemble-mean projected changes are consistent with those of the native GCM runoff. Specifically, this study addresses four questions: (1) how well the dataset reproduces global discharge patterns; (2) how much added value it provides relative to both the same modelling chain driven by

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



raw CMIP6 forcing and the native runoff output of the driving GCMs; (3) how consistent the nine ensemble members are with one another relative to the raw CMIP6 ensemble; and (4) whether the ensemble-mean projected changes of the derived runoff are consistent with those of the native GCM runoff.

By applying the same calibrated VIC configuration to all GCMs, GDDRP-CMIP6 reduces the confounding influence of heterogeneous native land-surface schemes and enables more direct comparison of runoff responses across GCMs and regions. The dataset provides daily total runoff from 1950 to 2100 under four SSPs at 0.25° and includes many islands that are absent from the coarse land–sea masks of the native GCM runoff because its land grid is defined independently of the GCM masks. Its routed discharge is evaluated globally against observations from 2,732 GRDC stations, and the regional performance and the main process limitations are documented. GDDRP-CMIP6 therefore provides an observation-evaluated, analysis-ready basis for spatially consistent assessments of floods, droughts, seasonal water availability, and long-term river-discharge changes, without requiring users to bias-correct and downscale climate forcing or to conduct global hydrological simulations themselves.

2 Data and methods

GDDRP-CMIP6 is produced by driving the Variable Infiltration Capacity (VIC) model with the daily 0.25° bias-corrected and downscaled atmospheric forcing of nine NEX-GDDP-CMIP6 members under the historical experiment and all SSP scenarios. A single VIC configuration with globally calibrated parameters is used for all members and scenarios. The product is the daily total runoff, the sum of the surface runoff and baseflow simulated by VIC, from 1950 to 2100, and the separate surface runoff and baseflow fields are available on request.

The product is validated by routing the simulated runoff through the CaMa-Flood global hydrodynamic model and comparing the river discharge characteristics with GRDC observations over the historical period. Two baseline runoff sets, the native runoff of the GCMs and the VIC runoff driven by the raw CMIP6 forcing, are routed in the same way to quantify the improvement of the product and to attribute it to the forcing correction and the hydrological model. The three runoff sets are further compared in terms of the inter-member consistency in the historical period and the projected change signals under SSP5-8.5.

2.1 NEX-GDDP-CMIP6 climatic forcing data

This study uses the NEX-GDDP-CMIP6 (Thrasher et al., 2022) dataset as the meteorological forcing for the global hydrological model. Produced by the NASA Earth Exchange team from the outputs of 35 CMIP6 GCMs, the dataset spans 1950–2100 at a spatial resolution of 0.25° (approximately 25 km) and a daily time step. The period 1950–2014 corresponds to the CMIP6 historical experiment, and 2015–2100 to four Shared Socioeconomic Pathway scenarios from ScenarioMIP, namely SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. To ensure uniform inputs to the hydrological model, we use the

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



eight variables common to all GCMs: daily mean temperature, daily maximum and minimum temperature, precipitation, relative humidity, longwave and shortwave radiation, and surface wind speed.

NEX-GDDP-CMIP6 is statistically downscaled with a daily variant of the Bias Correction and Spatial Disaggregation method (Thrasher et al., 2022; Wood et al., 2004). To maintain consistency across successive NEX-GDDP releases, the Global Meteorological Forcing Dataset (GMFD; Sheffield et al., 2006), which blends reanalysis with station observations, is retained as the correction reference. Bias correction uses 1960–2014 as the common reference period. The 0.25° daily GMFD fields are first interpolated to the native grid of each CMIP6 model. For each daily value to be corrected, candidate values are pooled within a ±15-day moving window across the reference period, from which empirical cumulative distribution functions are constructed for the observations and for the model historical experiment; each model value is then mapped from its quantile in the model distribution to the same quantile in the observed distribution. The corrected coarse-grid fields are subsequently disaggregated in space. Daily climatology is built from the GMFD reference-period data and smoothed by retaining the first three harmonics of a Fourier transform. After this climatology is interpolated to the target GCM grid, residual fields containing only the relative change are isolated from the corrected coarse-grid daily values, as differences for temperature variables and as ratios for the remaining variables. The residuals are bilinearly interpolated to the 0.25° grid and then added to, or multiplied by, the 0.25° fine-scale climatology. In the final product, the fine spatial structure is therefore supplied by the observed climatology, whereas the daily-to-interannual variability derives from the bias-corrected anomalies of each GCM.

For the future scenarios, BCSD cannot apply quantile mapping against an observed distribution as in the historical period, because no corresponding observations exist beyond the reference period. NEX-GDDP-CMIP6 therefore adopts a signal-preserving strategy. The long-term climate-change trend is first separated from each model's future time series by computing, for each calendar month independently, a 9-year running mean that characterizes the slowly varying trend of that month. Once this trend component is removed, the remaining high-frequency component falls within a range comparable to the reference period, so the quantile-mapping relationships established for the historical period can be applied; after correction, the previously separated trend is added back. In this way, the climate-change signal projected by each model is preserved in the corrected series: the bias correction acts only on the statistical distribution relative to the reference period and does not alter the magnitude of the model-projected long-term change.

Constrained by the computational cost of global 0.25° daily land-surface simulation, we select nine ensemble members from the NEX-GDDP-CMIP6 archive (Table 1), each from a different modelling center. All models use the r1i1p1f1 realization, and the scenarios comprise the historical experiment and SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, except for GFDL-CM4, which is available only under SSP2-4.5 and SSP5-8.5. The transient climate response (TCR) is the global-mean surface warming at the time of CO_2 doubling relative to pre-industrial levels in an experiment with CO_2 concentration increasing at 1 % per year, and it reflects the century-scale transient warming response, including the lag imposed by ocean heat uptake (Meehl et al., 2020). Synthesizing multiple lines of

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



evidence, the IPCC Sixth Assessment Report (AR6) assessed the very likely (5–95 %) range of TCR as 1.2–2.4 °C, with a central estimate of 1.8 °C (IPCC, 2021). The TCR of the nine selected models ranges from 1.33 °C (NorESM2-MM) to 2.32 °C (IPSL-CM6A-LR), covering most of the AR6 assessed range.

Table 1. Overview of the nine CMIP6 global climate models used in this study.

Models	Scenarios	TCR (°C)	Reference
ACCESS-CM2	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	2.1	(Bi et al., 2020)
EC-Earth3	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	2.3	(Döscher et al., 2022)
GFDL-CM4	Historical, SSP2-4.5 / 5-8.5	2.05	(Held et al., 2019)
INM-CM5-0	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	unknown	(Volodin et al., 2017)
IPSL-CM6A-LR	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	2.32	(Boucher et al., 2020)
MIROC6	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	1.55	(Tatebe et al., 2019)
MPI-ESM1-2-HR	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	1.66	(Müller et al., 2018)
MRI-ESM2-0	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	1.64	(Yukimoto et al., 2019)
NorESM2-MM	Historical, SSP1-2.6 / 2-4.5 / 3-7.0 / 5-8.5	1.33	(Seland et al., 2020)

2.2 Global hydrological modelling

This study employs the Variable Infiltration Capacity model (VIC; Liang et al., 1994), a macroscale land-surface hydrological model that represents the sub-grid heterogeneity of soil water storage capacity, vegetation, and elevation (Sect. S2 in the Supplement). VIC (version 4.2) is run in water-balance mode on a 0.25° grid at a daily time step to convert the meteorological forcing into surface runoff and baseflow. A common VIC configuration is applied to all GCM forcings to generate a structurally consistent runoff ensemble and facilitate comparison of hydrological responses across GCMs.

VIC is driven by the nine selected NEX-GDDP-CMIP6 ensemble members. Daily maximum and minimum temperature, precipitation, longwave and shortwave radiation, and wind speed are provided directly by the ensemble members, whereas actual vapor pressure and surface pressure are derived from relative humidity, temperature, and grid-mean elevation (Sect. S1 in the Supplement; Allan et al., 1998; Murray, 1967). Ten elevation bands are configured, with the mean elevation and areal fraction of each band within every 15-arcminute (0.25°) grid cell computed from the 1-arcminute MERIT DEM (Yamazaki et al., 2017); the precipitation partitioning coefficients are set equal to the areal fractions.

Soil parameters are taken from the globally calibrated parameter fields of Lin et al., (2019), in which the variable infiltration curve parameter (b), the thickness of the second soil layer (d2), and the fraction

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



of the maximum baseflow rate at which the nonlinear recession begins (Ds) were calibrated grid by grid using the Shuffled Complex Evolution algorithm (SCE-UA, Duan et al., 1992) and the Global Streamflow Characteristics Dataset (GSCD; Beck et al., 2015). Although the optimal parameter values are affected by biases in the meteorological forcing, calibration against independent data captures the relative spatial heterogeneity of soil responses (Sect. S2 in the Supplement). Compared with the default parameters, the calibrated parameters markedly reduce the underestimation of discharge, especially for small basins and low flows (Figs. S1 and S2 in the Supplement, using ACCESS-CM2). The main process limitation of this configuration is that glaciers and frozen soil are not represented, which affects the simulations in high-latitude and high-mountain regions.

2.3 Global hydrodynamic modelling

To evaluate the runoff product against observed discharge, the simulated total runoff (surface runoff plus baseflow) is routed into river discharge with the Catchment-based Macro-scale Floodplain model, CaMa-Flood (Yamazaki et al., 2011), a global river hydrodynamic model that discretizes the river network into unit-catchments and diagnoses water level and floodplain inundation from water storage and sub-grid topography. The river network and sub-grid topographic parameters are derived from MERIT Hydro (Yamazaki et al., 2019); the simulations in this study use a 1-arcminute high-resolution map and perform global simulations at a 6-arcminute (approximately 0.1°) resolution. River discharge is computed with the local inertial approximation of the shallow-water equations (Bates et al., 2010; Yamazaki et al., 2013), and a bifurcation scheme (Yamazaki et al., 2014) is included to represent flow in deltas and other multi-channel reaches (Sect. S3 in the Supplement). This study employs the GPU-accelerated version of CaMa-Flood (Kang et al., 2026), which reproduces the original model with negligible numerical differences and makes the global daily simulations of all members and scenarios computationally feasible. The same CaMa-Flood configuration is used for all runoff inputs in this study.

2.4 Experimental design

To quantify the improvement of the product relative to baseline simulations and to attribute its sources, this study constructs three parallel modelling chains that differ only in the runoff supplied to CaMa-Flood, while the river network, parameters, resolution, and simulation period of CaMa-Flood are kept identical. The first runoff set is taken directly from the native runoff generated by each GCM's own land-surface scheme, namely the total runoff variable, and is denoted Baseline 1. The second is generated by driving VIC with the raw (neither bias-corrected nor downscaled) CMIP6 meteorological variables, denoted Baseline 2. The third is the product of this study, generated by driving the same VIC configuration with the bias-corrected and downscaled NEX-GDDP-CMIP6 forcing. The three runoff sets are fed into identically configured CaMa-Flood simulations, yielding three sets of river discharge for the subsequent comparative evaluation. Pairwise comparisons among the three chains isolate the contributions of the individual components. Baseline 1, which uses native GCM runoff directly, represents one of the most common current practices, so the difference between GDDRP-CMIP6 and Baseline 1 reflects the overall improvement of the product. Baseline 2 differs from the

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



product only in meteorological forcing, so their difference reflects the gain contributed by the NEX-GDDP-CMIP6 bias correction and downscaling. The difference between Baseline 1 and Baseline 2 in turn reflects the contribution of replacing the GCMs' native runoff schemes with a global hydrological model.

2.5 Observed discharge data

The accuracy of the historical simulations is evaluated against daily observed discharge from the Global Runoff Data Centre (GRDC). Established under the auspices of the World Meteorological Organization, the GRDC provides quality-controlled river discharge observations from national hydrological agencies worldwide, together with station metadata including geographic coordinates, catchment area, and record periods.

To ensure observation quality and accurate spatial matching for model evaluation, stations were screened by the following criteria. First, to ensure reliable and continuous daily records, gaps in the observations were required to be no longer than 10 days; gaps of up to 10 days were filled by linear interpolation, and where a record contained longer interruptions, only the longer continuous segment was retained. Second, to obtain a robust station discharge climatology, the retained record had to exceed 10 years. Third, to prevent low flows and prolonged flow cessation, which lie beyond the model's simulation capability, from distorting the evaluation, stations were required to have a long-term mean daily discharge greater than 10 $m^3 s^{-1}$ and a Q90 greater than zero, where Q90 denotes the discharge corresponding to a 90 % exceedance probability. Fourth, the relative difference between the catchment area reported by the GRDC and the upstream area extracted by the routing model at the station location had to be within ±10 % to exclude stations with large geolocation errors. After screening, 2,732 stations were retained, spatially covering the world's major climate zones and representative land-surface types. Regions with sparse observations are concentrated mainly in relatively arid areas. In addition, station density is low in Southeast Asia, where most gauges were established too recently to meet the record-length criteria.

2.6 Evaluation against observed discharge

GCM historical simulations are free-running, and their internal variability is not synchronized in phase with the real world; a given date or year in the simulations therefore does not correspond to the same real event in the GRDC record, and the two time series cannot be compared directly. The simulated and observed discharge are therefore compared through their climatological statistics. For each station, the daily climatology, the multi-year mean discharge, and the flow quantiles at 90 %, 20 %, 10 %, 5 %, and 1 % exceedance probability (Q90, Q20, Q10, Q5, and Q1) are extracted from both the simulated and the observed discharge, and the ensemble medians of these characteristics across the nine members are used to compute the evaluation metrics.

Two metrics are used to evaluate these characteristics. The correlation coefficient of the daily climatology (CCC, Eq. (1)) measures how well the simulations reproduce the pattern and phase of the observed seasonal cycle, with higher values indicating closer agreement. The relative error (BIAS, Eq.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

(2)) measures the magnitude bias of the long-term water volume, represented by the multi-year mean discharge, of high flows, represented by Q10, and of low flows, represented by Q90. To compare the product with the baselines, the metric improvement (MI, Eq. (3)) is computed at each station; a positive MI indicates that the evaluated result is closer to the optimum than the baseline, and a negative MI indicates the opposite. To judge whether the improvement is robust at the regional scale, the Wilcoxon signed-rank test is applied to the MI values of all stations within each IPCC WGI reference region at the 95 % confidence level (Wilcoxon, 1945). This paired non-parametric test does not assume normality and is well suited to skewed hydrological metrics.

2.7 Ensemble consistency and future change signals

Each GCM provides one ensemble member. To characterize the inter-member spread in the historical period, a modified coefficient of variation (MCV, Eq. (4)) of the nine members is computed at each station for the long-term mean discharge and for the mean of Q1, Q5, Q10, and Q20, which represents high-flow behavior. Because the mean of biases may approach zero when positive and negative values cancel, the conventional coefficient of variation can become highly sensitive or even undefined; the root mean square is therefore used instead of the mean in the denominator.

Finally, to examine whether the ensemble-mean projected changes of the product are consistent with those of the driving GCMs, the change ratio (Eq. (5)) of the long-term mean discharge and Q1 between the SSP5-8.5 (2015–2100) and historical (1950–2014) periods is computed for each member, and the ensemble mean and the inter-member range of the change ratios are compared with those of Baseline 1. The same analysis is repeated for Baseline 2, which differs from the product only in the forcing correction, to separate the effect of the forcing correction from that of the common hydrological model. The comparison focuses on the ensemble-mean change rather than the change signal of each individual GCM, because replacing the native land-surface schemes of the GCMs with a common hydrological model is expected to modify the responses of individual members.

$$CCC = \frac{\sum_{i=1}^{n}(S_i-\bar{S})(O_i-\bar{O})}{\sqrt{\sum_{i=1}^{n}(S_i-\bar{S})^2}\sqrt{\sum_{i=1}^{n}(O_i-\bar{O})^2}} \quad (1)$$

$$BIAS = \frac{\sum_{i=1}^{n}(S_i-O_i)}{\sum_{i=1}^{n}O_i} \times 100\% \quad (2)$$

$$MI = |opt - m_{ref}| - |opt - m_{eval}| \quad (3)$$

$$MCV = \frac{\sqrt{\frac{\sum_{i=1}^{k}(BIAS_i-\overline{BIAS})^2}{k-1}}}{\sqrt{\frac{1}{k}\sum_{i=1}^{k}BIAS_i^2}} \quad (4)$$

$$ratio = \frac{I_{ssp585}}{I_{historical}} \quad (5)$$

where S_i and O_i denote the simulated and reference values at the ith sample, respectively; $\bar{S}$ and $\bar{O}$ are their arithmetic means over the n paired samples; m_{eval} and m_{ref} denote the metric values

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



of the evaluated and reference chains; opt is the optimal value of specific metrics, with 1 for CCC and 0 for *BIAS* respectively; for bias-type metrics, whose valid range is asymmetric (underestimation cannot exceed −100 %), a logarithmic transformation ($m = |ln(1 + bias/100)|$) is applied to render positive and negative biases numerically symmetric; k is the number of ensemble members; $I_{historical}$ and I_{ssp585} are the discharge characteristics under historical and SSP5-8.5 scenarios.

3 Dataset description

The dataset is archived at the Science Data Bank (ScienceDB) and, during the review process, is accessible at https://www.scidb.cn/en/anonymous/WkJiaUFq. It covers the nine NEX-GDDP-CMIP6 ensemble members and comprises the historical period (1950–2014) and all available SSP scenarios listed in Table 1 (2015–2100), at a spatial resolution of 0.25° and a daily time step, with a spatial extent of 180°W–180°E and 90°N–60°S. The land grid is defined independently of the GCM land–sea masks, so many small islands that are missing from the native GCM runoff are included. The publicly released variable is total runoff, the sum of surface runoff and baseflow, stored as one file per year for each GCM and scenario. Runoff is given in mm d^{-1}, and the archive contains 3509 files of about 120-150 MB each. The complete dataset has a total volume of about 500 GB. The dataset is distributed under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. Surface runoff and baseflow are available separately on request.

The data are stored in self-describing netCDF format, with file names following the structure: [product short name]_[GCM name]_[scenario]_[ensemble label]_[variable]_[yyyy].nc. Here, "[product short name]" is the unified identifier of this dataset, GDDRP-CMIP6; "[GCM name]" denotes the driving climate model; "[scenario]" is either historical or [SSP scenario name]; "[ensemble label]" identifies the ensemble member of the model (r1i1p1f1 for all models); "[variable]" is the variable name, TotalRunoff in the released dataset; and "[yyyy]" is the year covered by the file. For example, GDDRP-CMIP6_ACCESS-CM2_historical_r1i1p1f1_TotalRunoff_1950.nc contains the total runoff driven by ACCESS-CM2 under the historical experiment for 1950.

Each file contains one data variable, TotalRunoff (long_name "Total runoff (runoff + baseflow)", standard_name runoff_flux, units mm day-1), stored as 32-bit floating point with dimensions (time, lat, lon). The coordinate variables lat and lon (64-bit floating point, units degrees_north and degrees_east) give the centers of the 0.25° grid cells, from 89.875° N to 59.875° S (600 rows, ordered from north to south) and from 179.875° W to 179.875° E (1440 columns). The coordinate variable time (32-bit integer, units "days since 1950-01-01 00:00:00") has one value per day, marking the start of each day. The calendar attribute follows the driving GCM as provided by NEX-GDDP-CMIP6, so the number of time steps per file is 365 or 366 depending on the year and calendar. Ocean cells and cells outside the VIC land domain are set to the fill value (_FillValue = NaN), and the same land mask is used for all time steps, members and scenarios. The global attributes title, history and source record the content, creation time and source files of each file. The files are written in netCDF-4 format with the shuffle and deflate (level 4) filters. Before compression, runoff values are rounded to multiples of 1/1024 mm day-1 (about 0.001 mm day-1), as stated in the comment attribute.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Basic quality control was applied to the archive. All GCM, scenario and year combinations listed in Table 1 are present (3509 files), and each file contains the complete daily record of its year. Runoff values over land are finite and non-negative, and the land mask is identical in all time steps. The physical plausibility of the simulated runoff is assessed through the evaluation of routed discharge against GRDC observations in Sect. 4.

4 Technical validation

A runoff projection dataset intended for comparative climate-impact assessment should reproduce major observed discharge characteristics, provide a consistent historical reference across GCMs, and produce ensemble-mean projected changes broadly consistent with those of the driving models. The technical validation of GDDRP-CMIP6 therefore addresses these three requirements. First, historical performance is evaluated against observed discharge, and comparisons with two baseline simulations are used to assess the respective effects of bias-corrected forcing and the common VIC configuration. Second, inter-member consistency in the historical period is examined to determine whether the dataset provides a coherent basis for multi-model analysis. Third, inter-member spread and ensemble-mean projected changes are compared with the baselines to assess how the common modelling framework affects consistency and whether the ensemble-mean projected changes remain broadly consistent with those of the native GCM runoff. Together, these analyses characterize the observational performance, structural consistency, and suitability of GDDRP-CMIP6 for comparative climate-impact assessment.

4.1 Observed performance and sources of improvement

Figure 1 presents four evaluation metrics for the ensemble-median discharge simulated from GDDRP-CMIP6 against GRDC observations at global stations: the correlation coefficient of the daily climatology (CCC), the long-term mean bias (BIAS), the high-flow bias (Q10 BIAS), and the low-flow bias (Q90 BIAS). Below each map, boxplots summarize the metric globally and for four basin-area groups ($\leqslant 1$, 1–10, 10–100, and $>100 \times 10^4$ km^2), characterizing how each metric varies with basin scale. Table 2 lists the median of each metric for each group.

Table 2. Metric medians of global scale and different basin sizes

Basin scale (10^4 km^2)	Correlation	BIAS (%)	Q10 BIAS (%)	Q90 BIAS (%)
Global	0.83	-5.76	-10.18	10.65
≤1	0.82	-13.14	-16.43	-0.37
1-10	0.84	-1.91	-8.53	13.99
10-100	0.87	18.45	11.76	37.93
>100	0.88	18.77	11.19	37.91

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

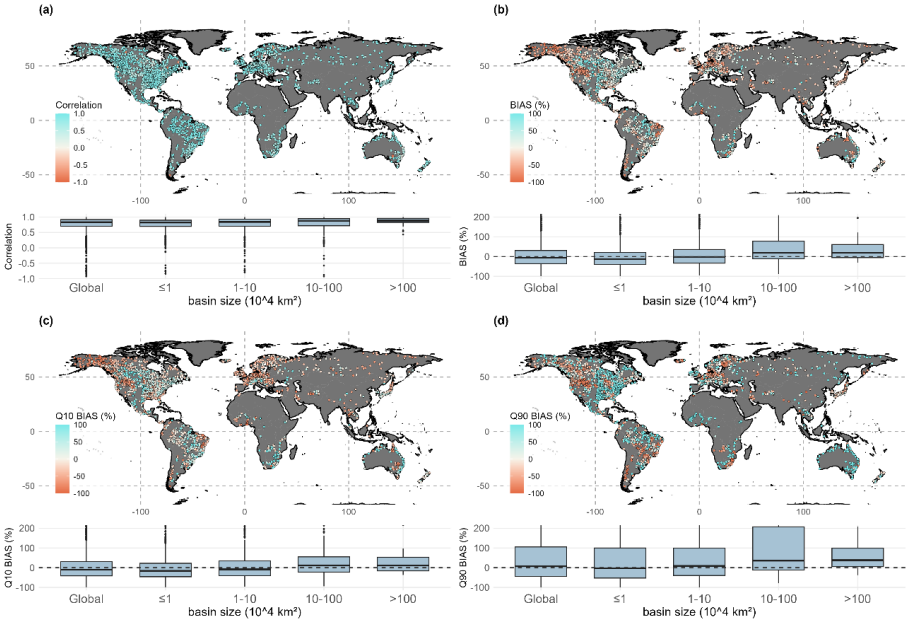


Figure 1. Evaluation of ensemble-median discharge characteristics simulated from GDDRP-CMIP6 against GRDC observations, including (a) the correlation coefficient of the daily climatology, (b) the long-term mean bias, (c) the high-flow bias, and (d) the low-flow bias. Boxplots beneath each map summarize the metric for all stations globally and for four basin-area groups.

The results show that the discharge simulated from GDDRP-CMIP6 generally reproduces the principal flow characteristics at most stations. CCC is high at stations worldwide, indicating that the seasonal cycle is well captured. The global median reaches 0.83, with performance improving and variability decreasing across stations as basin scale increases. The global median biases of the long-term mean, high flows, and low flows are −5.76 %, −10.18 %, and +10.65 %, respectively, indicating no severe systematic bias. Spatially, most stations attain small BIAS, with relatively pronounced underestimation confined to Western & Central Europe and the northwestern coastal region of North America. Regarding the reproduction of extremes, Q90 BIAS varies over a markedly wider range than Q10 BIAS, mainly because the bias calculation is more sensitive to low values. All three bias metrics tend to shift from underestimation in small basins to overestimation in large basins, with group medians ranging from −13.14 % to 18.77 %, −16.43 % to 11.19 %, and −0.37 % to 37.93 %, respectively.

To attribute the improvement of the product, its performance is compared with that of the two baselines over the IPCC WGI Reference Regions (Iturbide et al., 2020). Only regions containing more than five stations are analyzed, yielding 37 of the 46 regions. Figure 2 shows the change in the regional median of each metric, with significant changes outlined in black, and Table 3 classifies the significant cases by the signs of the changes relative to the two baselines. Relative to Baseline 1 (Fig. 2a, c, e, g),

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



GDDRP-CMIP6 improves in 79 of the 102 significant cases across the four metrics. The improvement is most evident for the seasonal cycle, with CCC increasing by at least 0.3 in all 30 significant regions (Fig. 2a), and for low flows, which improve in 21 of 29 significant regions (Fig. 2g). Deteriorations are confined to the bias metrics, with eight cases each for the long-term mean bias and low flows and seven for high flows.

The comparison with Baseline 2 (Fig. 2b, d, f, h) isolates the effect of the forcing correction, that is, the bias correction and downscaling of the forcing. The forcing correction significantly improves performance in 64 cases and degrades it in 36 (Table 3). Its benefit is clearest for CCC, which improves in 16 regions and deteriorates only in four regions in North America, and for high flows, with 18 improvements against nine deteriorations located mostly at northern high latitudes. For the long-term mean bias and low flows, positive effects still outnumber negative ones (16 versus 12 and 14 versus 11, respectively), with a particularly pronounced improvement of the mean bias over southern Africa. In 24 cases, the effect of the forcing correction is offset by that of VIC and leaves no significant change relative to Baseline 1; these cases occur mostly for the long-term mean bias (12) and high flows (8).

The remaining differences are attributable to replacing the native land-surface schemes with VIC. In 26 cases, the change relative to Baseline 1 is significant whereas that relative to Baseline 2 is not, so the product performs similarly to Baseline 2 and the change originates from VIC. Of these, 24 are improvements, 11 of which are for CCC, and the two deteriorations are low flows in eastern and northern Australia. In a further 21 cases, the effect of VIC outweighs a forcing effect of the opposite sign. It turns a negative forcing effect into an overall improvement in 13 cases, such as CCC in North America and the long-term mean bias in East Asia and western North America, and a positive forcing effect into an overall deterioration in eight cases, half of them for the long-term mean bias. The underestimation over Western & Central Europe and northwestern coastal North America stems partly from failures of the forcing correction and partly from the omission of glaciers and frozen soil in VIC. Overall, the forcing correction and VIC reinforce or compensate each other across regions, and the forcing correction is beneficial in nearly twice as many cases as it is detrimental.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.

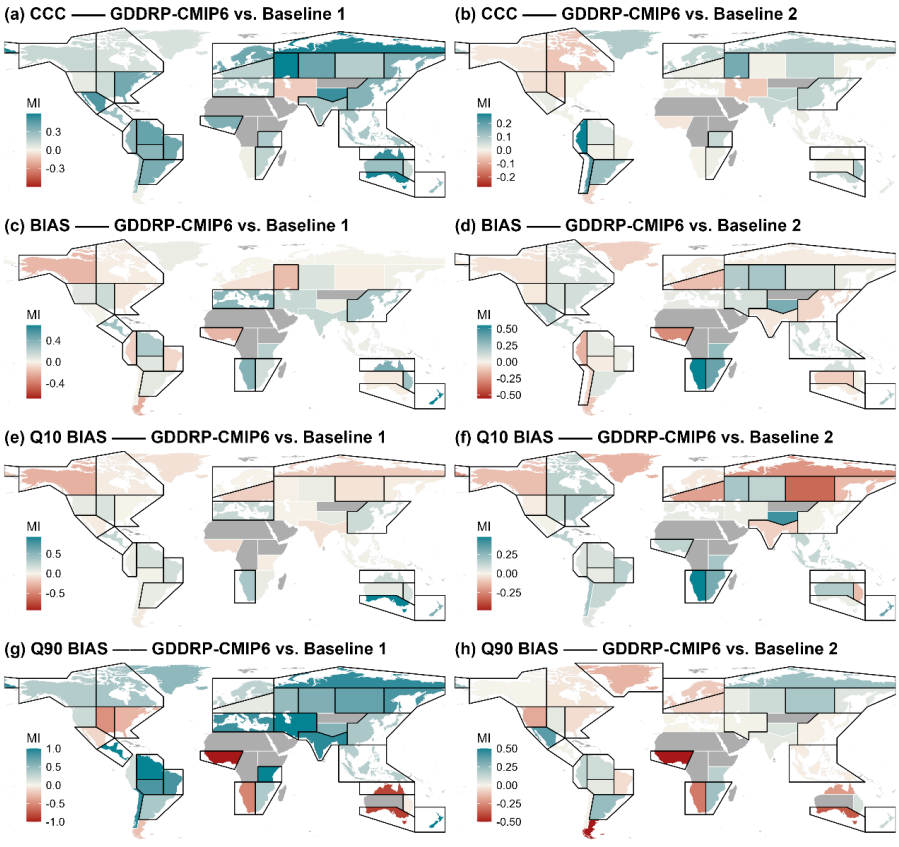


Figure 2. Regional median change of metric improvement under different comparison scenarios. (a, b) for the correlation coefficient of the daily climatology (CCC); (c, d) for the long-term mean bias (BIAS); (e, f) for the high-flow bias (Q10 BIAS); (g, h) for the low-flow bias (Q90 BIAS). The left column (a, c, e, g) compares GDDRP-CMIP6 with Baseline 1, and the right column (b, d, f, h) compares GDDRP-CMIP6 with Baseline 2; regions outlined in black show changes significant at the 95 % confidence level (Wilcoxon signed-rank test).

Table 3. Number of IPCC reference regions in each attribution category. "+" and "−" denote significant improvement and deterioration of GDDRP-CMIP6 relative to each baseline (Wilcoxon signed-rank test, 95 % confidence level), and "n.s." denotes no significant change. Q10 BIAS and Q90 BIAS refer to high and low flows, respectively.

vs Baseline 1	vs Baseline 2	Attribution	CCC	BIAS	Q10 BIAS	Q90 BIAS	Total
+	+	Forcing correction improves	15	6	10	11	42

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



−	−	Forcing correction degrades	0	4	5	4	13
+	−	VIC improves; forcing correction degrades	4	2	2	5	13
−	+	VIC degrades; forcing correction improves	0	4	2	2	8
+	n.s.	VIC improves	11	5	3	5	24
−	n.s.	VIC degrades	0	0	0	2	2
n.s.	+	Forcing correction improves; offset by VIC	1	6	6	1	14
n.s.	−	Forcing correction degrades; offset by VIC	0	6	2	2	10
Significant relative to Baseline 1			30	21	22	29	102

4.2 Consistency of historical simulation

Beyond the overall performance of the ensemble median, the spread of the individual ensemble members during the historical period also requires attention. In climate-change impact assessment, the ensemble mean or median is commonly taken as the representative result, on the ground that averaging smooths internal-variability noise and offsets model structural and parametric differences (Knutti et al., 2010; Tebaldi and Knutti, 2007), thereby improving the signal-to-noise ratio of the climate-change signal. However, the validity of this practice rests on the premise that the members fluctuate around a common state. When the members are already widely dispersed in the historical period, the ensemble median becomes a statistical composite of mutually incompatible hydrological states; for example, it may combine the seasonal timing of one member with the flow magnitude of another. Such a composite does not correspond to the response of any physically realizable system, so it loses its physical meaning, and results derived from it can be severely distorted. Moreover, because the models simulate the same historical period, the long-term statistics of their simulations should in principle converge towards the observational benchmark. Multi-model ensembles are widely regarded as an effective way to reduce the errors of individual models in runoff simulation (Hou et al., 2023). When the members are widely dispersed, the better performance of the ensemble median may arise largely from the cancellation of large, opposite-signed errors among members. An evaluation based on the ensemble median then masks the substantial differences in skill among models and may result in a wrong conclusion. This section therefore focuses on the inter-member spread of the long-term mean discharge, computing the MCV at each station.

Figure 3 presents the inter-member consistency of the discharge climatology, with global MCV distribution, cumulative distribution curves, and the climatology of each GCM from a representative station. The MCV of GDDRP-CMIP6 is systematically lower than that of Baseline 1 by about one order of magnitude. The median MCV is about 0.07 for GDDRP-CMIP6 and about 0.8 for Baseline 1, and about 80 % of the stations have an MCV below 0.2 in GDDRP-CMIP6, compared with about 5 % in Baseline 1. About half of the Baseline 1 stations have an MCV between 0.5 and 1.0, close to

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

the upper limit of about 1.06 that Eq. (4) allows for nine members, which is approached when the biases of the members have a near-zero mean and therefore cancel in the ensemble mean. The example station illustrates this difference: the nine curves driven by native GCM runoff differ markedly not only in their long-term means but also in the phase and magnitude of their seasonal climatologies. Spread of this magnitude means that the ensemble mean or median cannot adequately represent the individual members, and it also indicates that every model carries biases of differing severity. Under these circumstances, even where the ensemble-level accuracy of Baseline 1 appears acceptable at some stations, this does not imply that each GCM reproduces the historical characteristics at a comparable level. Moreover, the unequal biases across models further limit the usability of their absolute values. In flood-risk studies, hazards typically respond nonlinearly, with damage occurring once a threshold is exceeded, so judging whether the threshold is crossed depends on accurate absolute values.

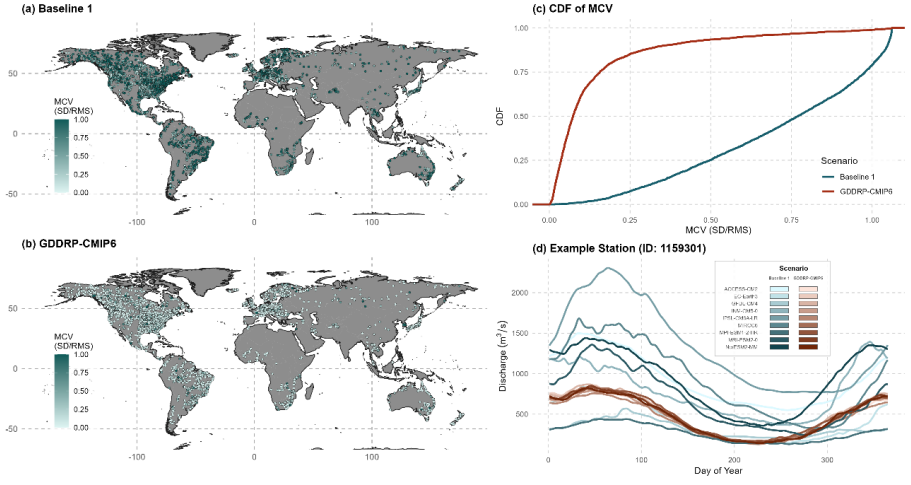


Figure 3. Comparison of MCV (Eq. (4)) of long-term mean discharge from Baseline 1 and GDDRP-CMIP6, including spatial distribution (a, b), cumulative distribution curves (c), and discharge climatologies of an example station (GRDC ID: 1159301).

In contrast to Baseline 1, the results driven by GDDRP-CMIP6 exhibit a high degree of inter-member consistency, with the climatologies of the members fluctuating around a common historical state, reflecting the removal of much of the GCM-inherent climatological bias by the forcing correction. Although this common state may still deviate from the observed one, the improved overall performance indicates that the product reproduces historical characteristics more accurately, and the more consistent results provide a firmer baseline for further runoff calibration and for climate-change impact assessment. Importantly, this gain in consistency extends to the extremes: Figure 4 adopts the same layout as Fig. 3 but targets high flows, with the example-station panel (d) showing the empirical frequency distributions of the two sets of annual maxima. The GDDRP-CMIP6 distributions form a tight cluster, in contrast to the wide scatter of Baseline 1. The median MCV of the high-flow metric is about 0.07 for GDDRP-CMIP6 and about 0.7 for Baseline 1 (Fig. 4c). Notably, the forcing correction

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

does not bring the members into complete agreement. The residual inter-member differences include internal variability and the differences among the driving GCMs that remain after bias correction. Overall, the GDDRP-CMIP6 members behave far more consistently in the historical period, the ensemble median becomes more representative, and the product can therefore support a broader range of applications.

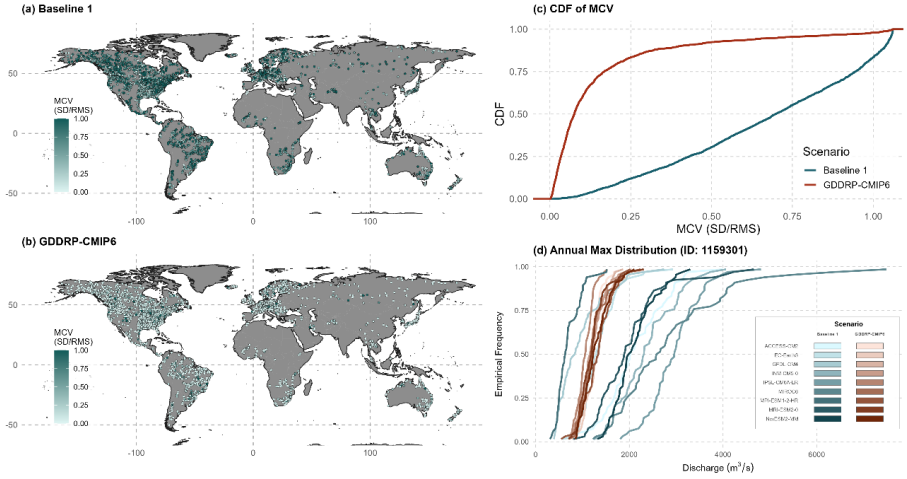


Figure 4. As Fig. 3 but for high flows, represented by the mean of Q1, Q5, Q10, and Q20. (d) is the frequency distribution of annual maxima.

4.3 Ensemble-mean projected changes

Owing to differences in climate response strength, the GCMs project different future changes. The spread of the ensemble projections is commonly used to represent future uncertainty and to bound the range of responses considered in decision-making. In this section, the change ratio of each member is computed at every station under SSP5-8.5 (2015–2100, relative to the historical period 1950–2014), the inter-member spread is measured as the range of the change ratios of the nine members, and the difference of the range between GDDRP-CMIP6 and Baseline 1 is further analyzed in terms of its spatial pattern and dependence on basin scale. Baseline 2 is analyzed in the same way to separate the contributions of the forcing correction and the common hydrological model.

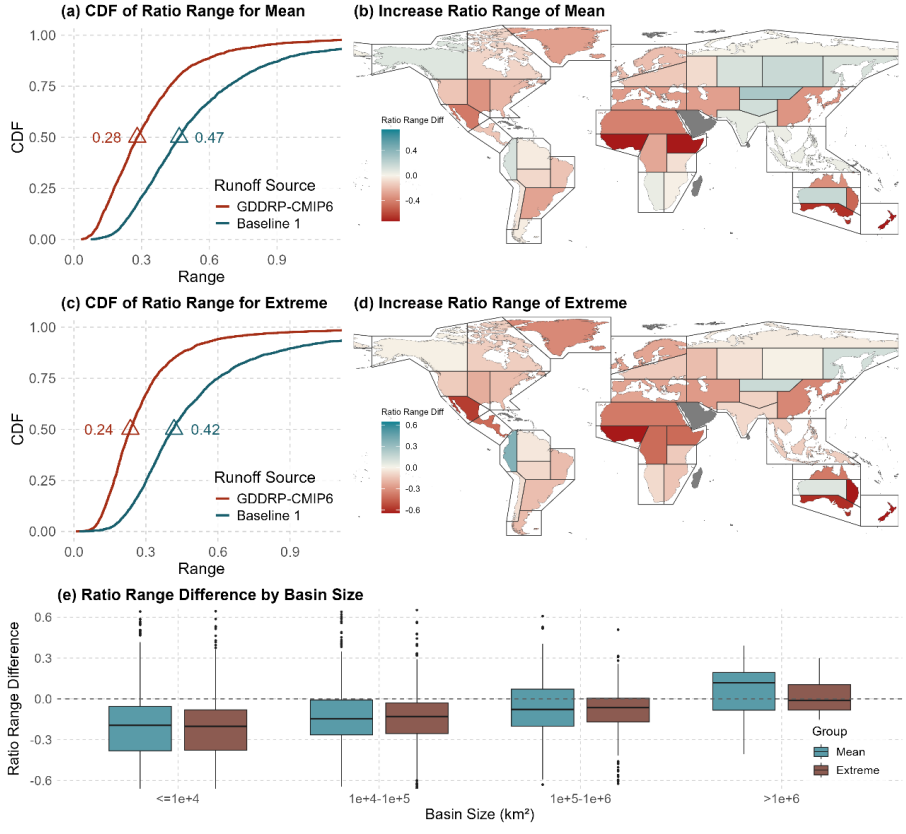
Figure 5 shows the ratio range differences. The left column (a, c) shows cumulative distributions of the ratio range for mean and extreme discharge, with triangles marking the medians; the right column (b, d) maps the difference in range (GDDRP-CMIP6 minus Baseline 1), where negative values denote reduced inter-member spread; and panel (e) groups this difference by basin area. The spread narrows almost everywhere: the global median range falls from 0.47 to 0.28 for mean discharge and from 0.42 to 0.24 for the extremes. The similar magnitude of the reduction for the two quantities indicates a systematic contraction rather than an incidental gain in one flow characteristic. The contraction is

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

519 attributable to the construction of the modelling chain. To separate the contributions of its components,
520 the same analysis was repeated for Baseline 2, in which VIC is driven by the raw CMIP6 forcing. The
521 inter-member range of Baseline 2 under SSP5-8.5 is close to that of GDDRP-CMIP6 (Fig. S3 in the
522 Supplement), indicating that the contraction arises mainly from the common hydrological model.
523 Because the land-surface processes of all members are represented by a single hydrological model,
524 the contribution of the differences among the native land-surface schemes of the GCMs to the inter-
525 member spread is removed. The same reasoning explains why the reduction increases as basin area
526 decreases: in small basins, local land-surface processes exert a more direct influence on discharge, so
527 unifying the hydrological scheme yields a larger reduction.



528

529 **Figure 5.** Inter-member spread of projected discharge changes under SSP5-8.5 for GDDRP-CMIP6
530 and Baseline 1. Cumulative distributions of the nine-member range of change ratios for mean (a) and
531 extreme (c) discharge; regional distribution of the difference in the range between the two schemes
532 for mean (b) and extreme (d) discharge; boxplots of the range difference grouped by basin area (e).

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

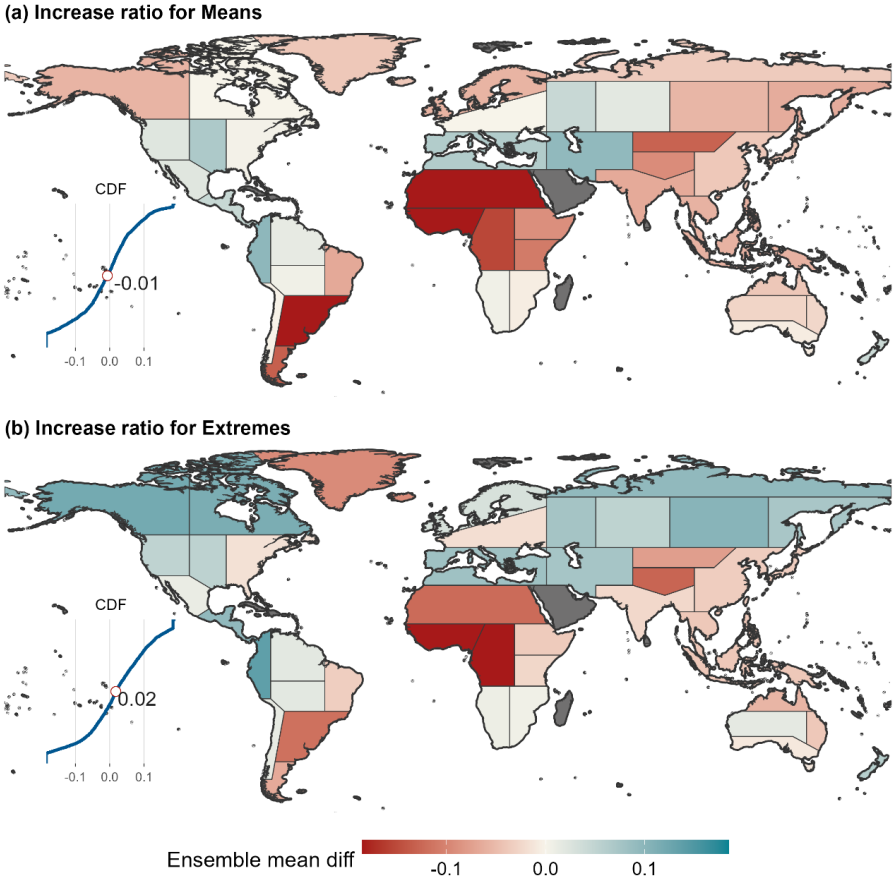


Figure 6. Difference in the ensemble-mean projected change between GDDRP-CMIP6 and Baseline 1 under SSP5-8.5. Insets show the cumulative distributions of the regional differences and their medians.

Because the native land-surface schemes are replaced by a common hydrological model, the responses of individual members are expected to differ from those of the raw GCMs, and the product is not intended to reproduce the change signal of each GCM. Instead, impact assessments commonly rely on the ensemble mean, so the product aims to produce ensemble-mean projected changes broadly consistent with those of the native GCM runoff. To test this consistency, Fig. 6 maps the regional difference between the two schemes' ensemble-mean change ratios for mean discharge (a) and extreme discharge (b), with insets in the lower-left corners showing the cumulative distribution of the differences and their medians. The global median difference in the ensemble-mean change ratio is only −0.01 for mean discharge and +0.02 for extreme discharge, and most of the regions fall within ±0.1, indicating that while the inter-member range narrows, GDDRP-CMIP6 produces ensemble-mean

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



projected changes broadly consistent with those of the native GCM runoff. Spatially, we focus on regions where the mean and the extremes show the same type of error (both overestimated or both underestimated) with significant differences, since such shifts are more likely to arise from systematic factors than from sampling noise. After excluding regions that lack representativeness owing to sparse stations, such as the Sahara and central Africa, the regions of concern are concentrated in southeastern South America and western Africa, both of which tend to underestimate the response magnitude. This may indicate that the global hydrological model introduces regional systematic biases there, and trend applications in these regions should be accompanied by cross-checks against the raw signals of the driving GCMs. Overall, GDDRP-CMIP6 narrows the inter-member spread of projected changes while producing ensemble-mean projected changes broadly consistent with those of the native GCM runoff.

5 Data availability

GDDRP-CMIP6 is archived at the Science Data Bank (ScienceDB) under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. During the review process, the dataset can be accessed anonymously through the ScienceDB review link:

https://www.scidb.cn/en/anonymous/WkJiaUFq

The NEX-GDDP-CMIP6 forcing is available from the NASA Center for Climate Simulation (https://doi.org/10.7917/OFSG3345; Thrasher et al., 2022). The GRDC discharge observations are available from the Global Runoff Data Centre (https://grdc.bafg.de). The native CMIP6 runoff and atmospheric variables used for the baselines are available from the Earth System Grid Federation. MERIT Hydro is available at http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/ (Yamazaki et al., 2019).

6 Code availability

The VIC model (version 4.2) is available at https://github.com/UW-Hydro/VIC. CaMa-Flood is available at https://github.com/global-hydrodynamics/CaMa-Flood_v4, and the GPU-accelerated version used in this study (Kang et al., 2026) is available at https://github.com/Kshy0/CaMa-Flood-GPU.

7 Conclusions

GDDRP-CMIP6 provides global daily total runoff at 0.25° from 1950 to 2100 for nine CMIP6 models under the historical experiment and four SSP scenarios. It is produced by driving a calibrated VIC configuration with the bias-corrected and downscaled NEX-GDDP-CMIP6 forcing, and includes islands absent from the native GCM land masks.

Evaluated at 2,732 GRDC stations after routing with CaMa-Flood, the derived discharge reproduces

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



the observed seasonal cycle (global median correlation 0.83) with moderate median biases in mean, high and low flows (−5.76 %, −10.18 % and +10.65 %). Relative to the native GCM runoff, performance improves in 79 of the 102 significant regional cases, most clearly for the seasonal cycle and low flows, and the forcing correction is beneficial in nearly twice as many cases as it is detrimental.

The common framework also makes the ensemble more consistent. The median MCV of the historical mean discharge falls from about 0.8 to about 0.07, with a similar reduction for high flows. Under SSP5-8.5, the median inter-member range of change ratios narrows from 0.47 to 0.28 for mean discharge and from 0.42 to 0.24 for extremes, mainly owing to the common hydrological model, while the global median difference in the ensemble-mean change ratio relative to the native GCM runoff is only −0.01 and +0.02, respectively.

Glaciers and frozen soil are not represented, which affects high-latitude and high-mountain regions, and the free-running historical simulations should be compared with observations only through climatological statistics. Trends in southeastern South America and western Africa should be cross-checked against the driving GCMs. Within these limits, GDDRP-CMIP6 provides an analysis-ready basis for comparable assessments of floods, droughts and water availability across regions and spatial scales.

Author contributions

Jiachao Chen: conceptualization, methodology, software, formal analysis, visualization, writing (original draft); Dai Yamazaki: supervision, writing (review and editing), Ming Pan: resources, writing (review and editing).

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank the Science Data Bank (ScienceDB) for hosting the dataset. This study is supported by CSTI SIP3 Program (Resilient Smart Network: JPJ012289) and MEXT “Program for the Advanced Studies of Climate Change Projection” (SENTAN: JPMXD0722680395).

References

Abeysingha, N., Ray, R., Tarkegn, T., 2025. Future hydro-climate extremes in the cypress creek watershed in Texas under different CMIP6 scenarios. Sustain. WATER Resour. Manag. 11. https://doi.org/10.1007/s40899-025-01285-6

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



609 Allan, R., Pereira, L., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water
610 requirements-FAO Irrigation and drainage paper 56.
611 Baogang, Y., Linxiao, W., Hongyu, T., Yonghua, L., Yong, W., Fen, Z., Jie, Z., Tianyu, Z., Tananbang,
612 L., 2024. Future changes in extremes across China based on NEX-GDDP-CMIP6 models.
613 Clim. Dyn. 62, 9587–9617. https://doi.org/10.1007/s00382-024-07408-7
614 Bass, B., Thackeray, C., Hall, A., Rahimi, S., Huang, L., 2024. A Novel Emergent Constraint Approach
615 for Refining Regional Climate Model Projections of Peak Flow Timing. Geophys. Res. Lett.
616 51. https://doi.org/10.1029/2024GL108575
617 Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water
618 equations for efficient two-dimensional flood inundation modelling. J. Hydrol. 387, 33–45.
619 https://doi.org/10.1016/j.jhydrol.2010.03.027
620 Beck, H.E., Roo, A. de, Dijk, A.I.J.M. van, 2015. Global Maps of Streamflow Characteristics Based
621 on Observations from Several Thousand Catchments. J. Hydrometeorol. 16, 1478–1501.
622 https://doi.org/10.1175/JHM-D-14-0155.1
623 Belay, H., Melesse, A., Tegegne, G., 2024. Evaluation and comparison of the performances of the
624 CMIP5 and CMIP6 models in reproducing extreme rainfall in the Upper Blue Nile basin of
625 Ethiopia. Theor. Appl. Climatol. 155, 9471–9496. https://doi.org/10.1007/s00704-024-
626 05187-z
627 Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky,
628 J., Rashid, H.A., Dobrohotoff, P., Mackallah, C., Yan, H., Hirst, A., Savita, A., Dias, F.B.,
629 Woodhouse, M., Fiedler, R., Heerdegen, A., 2020. Configuration and spin-up of ACCESS-
630 CM2, the new generation Australian Community Climate and Earth System Simulator
631 Coupled Model. J. South. Hemisphere Earth Syst. Sci. 70, 225–251.
632 https://doi.org/10.1071/ES19040
633 Bjarke, N., Livneh, B., Barsugli, J., Pendergrass, A., Small, E., 2024. Evaluating Large-Storm
634 Dominance in High-Resolution GCMs and Observations Across the Western Contiguous
635 United States. EARTHS FUTURE 12. https://doi.org/10.1029/2023EF004289
636 Boucher, O., Servonnat, J., Albright, A.L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet,
637 R., Bony, S., Bopp, L., 2020. Presentation and Evaluation of the IPSL-CM6A-LR Climate
638 Model. J. Adv. Model. Earth Syst. 12, e2019MS002010.
639 https://doi.org/10.1029/2019MS002010
640 Cannon, F., Carvalho, L., Jones, C., Norris, J., Bookhagen, B., Kiladis, G., 2017. Effects of
641 topographic smoothing on the simulation of winter precipitation in High Mountain Asia. J.
642 Geophys. Res.-Atmos. 122, 1456–1474. https://doi.org/10.1002/2016JD026038
643 Chanda, N., Chintalacheruvu, M., Choudhary, A., 2025. Hydrological modelling of a mountainous
644 watershed: simulating streamflows under present and projected climate conditions. Theor.
645 Appl. Climatol. 156. https://doi.org/10.1007/s00704-024-05250-9
646 Chandrasekar, A., Boeing, F., Thober, S., Samaniego, L., Rakovec, O., Marx, A., 2026. Increased high
647 flows in Germany under 1.5 K, 2.0 K and 3.0 K warming based on a large regional climate
648 model ensemble. Clim. CHANGE 179. https://doi.org/10.1007/s10584-026-04165-w
649 Cheng, Z., Zhou, T., Chen, X., Zhang, L., Qian, Y., Wang, Z., He, L., Leung, L., 2024. Understanding
650 the Biases in Global Monsoon Simulations from the Perspective of Atmospheric Energy

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



651 Transport. J. Clim. 37, 4647–4666. https://doi.org/10.1175/JCLI-D-23-0444.1
652 de Alvarenga, C., Alvarenga, L., Melo, P., Tomasella, J., Pinto, P., de Mello, C., Isidoro, J., 2026.
653 Impacts of Climate Change on the Hydrology of a Highly Disturbed Tropical River Basin.
654 EARTH 7. https://doi.org/10.3390/earth7020052
655 Dioha, E., Chung, E., Ayugi, B., Babaousmail, H., Sian, K., 2024. Quantifying the Added Value in the
656 NEX-GDDP-CMIP6 Models as Compared to Native CMIP6 in Simulating Africa's Diverse
657 Precipitation Climatology. EARTH Syst. Environ. 8, 417–436.
658 https://doi.org/10.1007/s41748-024-00397-x
659 Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R.,
660 Boussetta, S., Caron, L.-P., Carver, G., 2022. The EC-Earth3 Earth system model for the
661 Coupled Model Intercomparison Project 6. Geosci. Model Dev. 15, 2973–3020.
662 https://doi.org/10.5194/gmd-15-2973-2022
663 Du, Y., Wang, DG, Zhu, J., Wang, DY, Qi, X., Cai, J., 2022. Comprehensive assessment of CMIP5 and
664 CMIP6 models in simulating and projecting precipitation over the global land. Int. J. Climatol.
665 42, 6859–6875. https://doi.org/10.1002/joc.7616
666 Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual
667 rainfall-runoff models. Water Resour. Res. 28, 1015–1031.
668 https://doi.org/10.1029/91WR02985
669 Exbrayat, J., Buytaert, W., Timbe, E., Windhorst, D., Breuer, L., 2014. Addressing sources of
670 uncertainty in runoff projections for a data scarce catchment in the Ecuadorian Andes. Clim.
671 CHANGE 125, 221–235. https://doi.org/10.1007/s10584-014-1160-x
672 Galavi, H., Kamal, M., Mirzaei, M., Ebrahimian, M., 2019. Assessing the contribution of different
673 uncertainty sources in streamflow projections. Theor. Appl. Climatol. 137, 1289–1303.
674 https://doi.org/10.1007/s00704-018-2669-0
675 Giorgi, F., Gutowski, W., 2015. Regional Dynamical Downscaling and the CORDEX Initiative, in:
676 Gadgil, A., Tomich, T. (Eds.), ANNUAL REVIEW OF ENVIRONMENT AND
677 RESOURCES, VOL 40. pp. 467–490. https://doi.org/10.1146/annurev-environ-102014-
678 021217
679 Giorgi, F., Raffaele, F., 2022. On the dependency of GCM-based regional surface climate change
680 projections on model biases, resolution and climate sensitivity. Clim. Dyn. 58, 2843–2862.
681 https://doi.org/10.1007/s00382-021-06037-8
682 He, J., Yao, L., Zhang, Y., Chen, Y., Wang, D., 2026. Improving regional precipitation simulation over
683 South China: a comparative evaluation of NEX-GDDP-CMIP6 and CMIP6 models. Atmos.
684 Res. 336, 108822. https://doi.org/10.1016/j.atmosres.2026.108822
685 Held, I.M., Guo, H., Adcroft, A., Dunne, J.P., Horowitz, L.W., Krasting, J., Shevliakova, E., Winton,
686 M., Zhao, M., Bushuk, M., Wittenberg, A.T., Wyman, B., Xiang, B., Zhang, R., Anderson, W.,
687 Balaji, V., Donner, L., Dunne, K., Durachta, J., Gauthier, P.P.G., Ginoux, P., Golaz, J.-C.,
688 Griffies, S.M., Hallberg, R., Harris, L., Harrison, M., Hurlin, W., John, J., Lin, P., Lin, S.-J.,
689 Malyshev, S., Menzel, R., Milly, P.C.D., Ming, Y., Naik, V., Paynter, D., Paulot, F.,
690 Ramaswamy, V., Reichl, B., Robinson, T., Rosati, A., Seman, C., Silvers, L.G., Underwood,
691 S., Zadeh, N., 2019. Structure and Performance of GFDL's CM4.0 Climate Model. J. Adv.
692 Model. Earth Syst. 11, 3691–3727. https://doi.org/10.1029/2019MS001829

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



693 Hoang, L., Lauri, H., Kummu, M., Koponen, J., van Vliet, M., Supit, I., Leemans, R., Kabat, P.,
694 Ludwig, F., 2016. Mekong River flow and hydrological extremes under climate change.
695 Hydrol. EARTH Syst. Sci. 20, 3027–3041. https://doi.org/10.5194/hess-20-3027-2016
696 Hou, Y., Guo, H., Yang, Y., Liu, W., 2023. Global Evaluation of Runoff Simulation From Climate,
697 Hydrological and Land Surface Models. Water Resour. Res. 59, e2021WR031817.
698 https://doi.org/10.1029/2021WR031817
699 IPCC, 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to
700 the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge
701 University Press, Cambridge, United Kingdom and New York, NY, USA.
702 Iqbal, Z., Shahid, S., Ahmed, K., Ismail, T., Ziarh, G.F., Chung, E.-S., Wang, X., 2021. Evaluation of
703 CMIP6 GCM rainfall in mainland Southeast Asia. Atmos. Res. 254, 105525.
704 https://doi.org/10.1016/j.atmosres.2021.105525
705 Iturbide, M., Gutiérrez, J.M., Alves, L.M., Bedia, J., Cerezo-Mota, R., Cimadevilla, E., Cofiño, A.S.,
706 Di Luca, A., Faria, S.H., Gorodetskaya, I.V., Hauser, M., Herrera, S., Hennessy, K., Hewitt,
707 H.T., Jones, R.G., Krakovska, S., Manzanas, R., Martínez-Castro, D., Narisma, G.T., Nurhati,
708 I.S., Pinto, I., Seneviratne, S.I., van den Hurk, B., Vera, C.S., 2020. An update of IPCC climate
709 reference regions for subcontinental analysis of climate model data: definition and aggregated
710 datasets. Earth Syst. Sci. Data 12, 2959–2970. https://doi.org/10.5194/essd-12-2959-2020
711 Jain, S., Salunke, P., Mishra, S.K., Sahany, S., Choudhary, N., 2019. Advantage of NEX-GDDP over
712 CMIP5 and CORDEX Data: Indian Summer Monsoon. Atmos. Res. 228, 152–160.
713 https://doi.org/10.1016/j.atmosres.2019.05.026
714 Kang, S., Yin, J., Yamazaki, D., 2026. CaMa-Flood-GPU: a GPU-based hydrodynamic model
715 implementation for scalable global simulations. Geosci. Model Dev. 19, 5623–5640.
716 https://doi.org/10.5194/gmd-19-5623-2026
717 Kay, A., Davies, H., Bell, V., Jones, R., 2009. Comparison of uncertainty sources for climate change
718 impacts: flood frequency in England. Clim. CHANGE 92, 41–63.
719 https://doi.org/10.1007/s10584-008-9471-4
720 Kim, Y.-H., Min, S.-K., Zhang, X., Sillmann, J., Sandstad, M., 2020. Evaluation of the CMIP6 multi-
721 model ensemble for climate extreme indices. WEATHER Clim. Extrem. 29, 100269.
722 https://doi.org/10.1016/j.wace.2020.100269
723 Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G.A., 2010. Challenges in Combining
724 Projections from Multiple Climate Models. J. Clim. 23, 2739–2758.
725 https://doi.org/10.1175/2009JCLI3361.1
726 Kurniadi, A., Weller, E., Kim, Y., Min, S., 2023. Evaluation of Coupled Model Intercomparison Project
727 Phase 6 model-simulated extreme precipitation over Indonesia. Int. J. Climatol. 43, 174–196.
728 https://doi.org/10.1002/joc.7744
729 Lange, S., 2019. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD
730 (v1.0). Geosci. MODEL Dev. 12, 3055–3070. https://doi.org/10.5194/gmd-12-3055-2019
731 Lee, J., Kim, YO, Kim, Y, 2017. A new uncertainty analysis in the climate change impact assessment.
732 Int. J. Climatol. 37, 3837–3846. https://doi.org/10.1002/joc.4957
733 Liang, X., Letternmaier, D., Wood, E., Burges, S., 1994. A simple hydrologically based model of land
734 surface water and energy fluxes for general circulation models. J. Geophys. Res. Atmos. 99,

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



14415–14428. https://doi.org/10.1029/94JD00483

Lin, P., Pan, M., Beck, H.E., Yang, Y., Yamazaki, D., Frasson, R., David, C.H., Durand, M., Pavelsky, T.M., Allen, G.H., Gleason, C.J., Wood, E.F., 2019. Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches. Water Resour. Res. 55, 6499–6516. https://doi.org/10.1029/2019WR025287

Lun, Y., Liu, L., Cheng, L., Li, X., Li, H., Xu, Z., 2021. Assessment of GCMs simulation performance for precipitation and temperature from CMIP5 to CMIP6 over the Tibetan Plateau. Int. J. Climatol. 41, 3994–4018. https://doi.org/10.1002/joc.7055

Manikanta, V., Reddy, V., Das, J., 2024. Investigating the Limitations of Multi-Model Ensembling of Climate Model Outputs in Capturing Climate Extremes. Int. J. Climatol. 44, 5711–5726. https://doi.org/10.1002/joc.8660

Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themessl, M., Venema, V., Chun, K., Goodess, C., Jones, R., Onof, C., Vrac, M., Thiele-Eich, I., 2010. PRECIPITATION DOWNSCALING UNDER CLIMATE CHANGE: RECENT DEVELOPMENTS TO BRIDGE THE GAP BETWEEN DYNAMICAL MODELS AND THE END USER. Rev. Geophys. 48. https://doi.org/10.1029/2009RG000314

Maurya, S., Srivastava, P., Zhuo, L., Yaduvanshi, A., Mall, R., 2023. Future Climate Change Impact on the Streamflow of Mahi River Basin Under Different General Circulation Model Scenarios. WATER Resour. Manag. 37, 2675–2696. https://doi.org/10.1007/s11269-022-03372-1

Meehl, G.A., Senior, C.A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R.J., Taylor, K.E., Schlund, M., 2020. Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. Sci. Adv. 6, eaba1981. https://doi.org/10.1126/sciadv.aba1981

Müller, W.A., Jungclaus, J.H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., 2018. A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). J. Adv. Model. Earth Syst. 10, 1383–1413. https://doi.org/10.1029/2017MS001217

Murray, F.W., 1967. On the Computation of Saturation Vapor Pressure. J. Appl. Meteorol. Climatol. 6, 203–204.

Musie, M., Sen, S., Srivastava, P., 2020. Application of CORDEX-AFRICA and NEX-GDDP datasets for hydrologic projections under climate change in Lake Ziway sub-basin, Ethiopia. J. Hydrol. Reg. Stud. 31, 100721. https://doi.org/10.1016/j.ejrh.2020.100721

Muzava, M., Mhizha, A., Rwasoka, D.T., Gumindoga, W., 2025. Climate change impacts on the small-scale hydropower potential for the Pungwe B hydropower scheme in Zimbabwe using a multi-model climate ensemble. Phys. Chem. Earth Parts ABC 141, 104118. https://doi.org/10.1016/j.pce.2025.104118

Papadimitriou, L., Koutroulis, A., Grillakis, M., Tsanis, I., 2017. The effect of GCM biases on global runoff simulations of a land surface model. Hydrol. EARTH Syst. Sci. 21, 4379–4401. https://doi.org/10.5194/hess-21-4379-2017

Prudhomme, C., Reynard, N., Crooks, S., 2002. Downscaling of global climate models for flood frequency analysis: where are we now? Hydrol. Process. 16, 1137–1150.

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



https://doi.org/10.1002/hyp.1054

Roulo, D., Pichuka, S., Menon, R., 2026. From Global to Basin-Scale: Identifying Best-Performing CMIP6 Models Across Indian River Basins Through Downscaled Precipitation Products. Int. J. Climatol. https://doi.org/10.1002/joc.70367

Sahany, S., Mishra, S., Salunke, P., 2019. Historical simulations and climate change projections over India by NCAR CCSM4: CMIP5 vs. NEX-GDDP. Theor. Appl. Climatol. 135, 1423–1433. https://doi.org/10.1007/s00704-018-2455-z

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L.S., Debernard, J.B., Gupta, A.K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K.S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I.H.H., 2020. Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. Geosci. Model Dev. 13, 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. J. Clim. 19. https://doi.org/10.1175/JCLI3790.1

Singhal, P., Goel, N., Bronstert, A., Vormoor, K., Agarwal, A., Arora, M., 2025. Understanding the flood seasonality in a Himalayan River Basin under changing climate. Theor. Appl. Climatol. 156. https://doi.org/10.1007/s00704-025-05585-x

Sooraj, K.P., Terray, P., Aswale, A.M., 2026. Uncertainty in projected changes of Indian Summer Monsoon Rainfall by CMIP6 models. Atmos. Res. 328, 108424. https://doi.org/10.1016/j.atmosres.2025.108424

Sreedevi, S., Eldho, T.I., Jayasankar, T., 2022. Physically-based distributed modelling of the hydrology and soil erosion under changes in landuse and climate of a humid tropical river basin. CATENA 217, 106427. https://doi.org/10.1016/j.catena.2022.106427

Su, T., Liu, Z., Duan, Q., Mao, X., Xu, W., 2025. Advancing the Reliability of Future Hydrological Projections in a Snow-Dominated Alpine Watershed: Integrating Uncertainty Decomposition and CycleGAN Bias Correction. EARTHS FUTURE 13. https://doi.org/10.1029/2024EF005615

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., 2019. Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. Geosci. Model Dev. 12, 2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Tebaldi, C., Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 365, 2053–2075. https://doi.org/10.1098/rsta.2007.2076

Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., Nemani, R., 2022. NASA Global Daily Downscaled Projections, CMIP6. Sci. Data 9, 262. https://doi.org/10.1038/s41597-022-01393-4

Tian, J., Zhang, Z., Fu, Y., Tao, H., Zhu, B., Yang, L., 2023. Does CMIP6 inspire more confidence in projecting precipitation over China? Int. J. Climatol. 43, 4965–4986. https://doi.org/10.1002/joc.8127

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



819 Tong, Y., Gao, X., Xu, Y., Cui, X., Giorgi, F., 2024. Inter-Comparison of Precipitation Simulation and
820 Future Projections Over China From an Ensemble of Multi-GCM Driven RCM Simulations.
821 J. Geophys. Res.-Atmos. 129. https://doi.org/10.1029/2023JD040166
822 Usman, M., Manzanas, R., Ndehedehe, C., Ahmad, B., Adeyeri, O., Dudzai, C., 2022. On the Benefits
823 of Bias Correction Techniques for Streamflow Simulation in Complex Terrain Catchments: A
824 Case-Study for the Chitral River Basin in Pakistan. HYDROLOGY 9.
825 https://doi.org/10.3390/hydrology9110188
826 Volodin, E.M., Mortikov, E.V., Kostrykin, S.V., Galin, V.Ya., Lykossov, V.N., Gritsun, A.S., Diansky,
827 N.A., Gusev, A.V., Iakovlev, N.G., 2017. Simulation of the present-day climate with the
828 climate model INMCM5. Clim. Dyn. 49, 3715–3734. https://doi.org/10.1007/s00382-017-
829 3539-7
830 Wang, D., Liu, J., Wang, H., Shao, W., Mei, C., Ding, X., 2022. Performance evaluations of CMIP6
831 and CMIP5 models for precipitation simulation over the Hanjiang River Basin, China. J.
832 Water Clim. Change 13, 2089–2106. https://doi.org/10.2166/wcc.2022.402
833 Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. Biom. Bull. 1, 80–83.
834 https://doi.org/10.2307/3001968
835 Wood, A., Leung, L., Sridhar, V., Lettenmaier, D., 2004. Hydrologic implications of dynamical and
836 statistical approaches to downscaling climate model outputs. Clim. CHANGE 62, 189–216.
837 https://doi.org/10.1023/B:CLIM.0000013685.99609.9e
838 Yamazaki, D., de Almeida, G.A.M., Bates, P.D., 2013. Improving computational efficiency in global
839 river models by implementing the local inertial flow equation and a vector-based river
840 network map. Water Resour. Res. 49, 7221–7235. https://doi.org/10.1002/wrcr.20552
841 Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H., Pavelsky, T.M., 2019. MERIT Hydro:
842 A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. Water
843 Resour. Res. 55, 5053–5073. https://doi.org/10.1029/2019WR024873
844 Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C.,
845 Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. Geophys. Res.
846 Lett. 44, 5844–5853. https://doi.org/10.1002/2017GL072874
847 Yamazaki, D., Kanae, S., Kim, H., Oki, T., 2011. A physically based description of floodplain
848 inundation dynamics in a global river routing model. Water Resour. Res. 47.
849 https://doi.org/10.1029/2010WR009726
850 Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y., Bates, P.D., 2014. Regional flood dynamics in a
851 bifurcating mega delta simulated in a global river model. Geophys. Res. Lett. 41, 3127–3135.
852 https://doi.org/10.1002/2014GL059744
853 Yan, T., Bai, J., Arsenio, T., Liu, J., Shen, Z., 2019. Future climate change impacts on streamflow and
854 nitrogen exports based on CMIP5 projection in the Miyun Reservoir Basin, China. Ecohydrol.
855 Hydrobiol., New advancements in Soil and Water Assessment Tool (SWAT) for
856 ecohydrological modelling and application 19, 266–278.
857 https://doi.org/10.1016/j.ecohyd.2018.09.001
858 Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M.,
859 Tanaka, T.Y., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi,
860 Y., Ishii, M., 2019. The Meteorological Research Institute Earth System Model Version 2.0,

https://doi.org/10.5194/essd-2026-862
Preprint. Discussion started: 23 November 2026
© Author(s) 2026. CC BY 4.0 License.



MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. J. Meteorol. Soc. Jpn. Ser II 97, 931–965. https://doi.org/10.2151/jmsj.2019-051

Zam, P., Shrestha, S., Budhathoki, A., 2021. Assessment of climate change impact on hydrology of a transboundary river of Bhutan and India. J. Water Clim. Change 12, 3224–3239. https://doi.org/10.2166/wcc.2021.338

Zhang, S., Chen, J., Gu, L., 2022. Overall uncertainty of climate change impacts on watershed hydrology in China. Int. J. Climatol. 42, 507–520. https://doi.org/10.1002/joc.7257

Zhang, X., Hua, L., Jiang, D., 2022. Assessment of CMIP6 model performance for temperature and precipitation in Xinjiang, China. Atmos. Ocean. Sci. Lett. 15, 100128. https://doi.org/10.1016/j.aosl.2021.100128

Zhao, Y., Dong, N., Li, Z., Zhang, W., Yang, M., Wang, H., 2021. Future precipitation, hydrology and hydropower generation in the Yalong River Basin: Projections and analysis. J. Hydrol. 602, 126738. https://doi.org/10.1016/j.jhydrol.2021.126738