www.nature.com/scientificdata

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A ten-year (2012–2021) fine-resolution (1 km, hourly) precipitation dataset over southeastern Tibetan Plateau

Dayang Li[1,2,3], Ye Shen[1,3], Xianyi Ye[1,3], Xiaohua Fu[1,3,4], Yang Yang[5], Tinghai Ou[6], Deliang Chen[7], Fuqiang Tian[8] & Long Yang[1,3]✉

Intense sub-daily precipitation on the Tibetan Plateau (TP) can trigger a cascade of natural hazards (flash floods, debris flows, etc.), causing significant environmental impacts. Current precipitation products for the TP often lack sufficient spatial or temporal resolution, hindering accurate characterization and mitigation of precipitation-related hazards. Here, we carry out 10-year (2012–2021, June to September) convection-permitting model simulations based on the Weather Research and Forecasting (WRF) model over the TP. We adopt four one-way nested domains, with the innermost domain providing a 1-km, hourly resolution over the southeastern TP. Model performance is enhanced using a spectral nudging scheme. Our simulated precipitation data are compared against *in-situ* rain gauge observations and state-of-the-art gridded precipitation products (i.e., HAR v2, ERA5-Land, and IMERG) over the southeastern TP. Our precipitation dataset demonstrates superior accuracy in mean absolute error, root mean square error, and bias, compared to the other three products. It effectively captures the key feature of the diurnal precipitation cycle and the non-monotonic dependence of precipitation on complex topography over the southeastern TP.

Background & Summary

The Tibetan Plateau (TP) is often known as the "Asian Water Tower". It is the source of ten major rivers such as Yangtze, Yellow, Yarlung Zangbo (Brahmaputra), Salween, and Mekong (Lancang) that sustain nearly two billion people over the Asian continent[1–5]. TP is highly vulnerable to precipitation -related hazards that frequently lead to notable damages to property, infrastructure, and agriculture, as well as losses of human lives[6,7]. Intense rain storms on sub-daily timescales are crucial triggers of natural hazards, e.g., flash floods, debris flows, etc.[8]. Obtaining high-resolution precipitation data in both space and time is crucial to mitigate these hazards, but it remains challenging over the TP.

Precipitation characterization over the TP is commonly based on *in-situ* gauge observations, satellite retrievals, and reanalysis of climate model outputs[9]. However, inherent limitations and uncertainties exist in these products. For instance, the density of rain gauge networks is significantly below the minimum recommended by the World Meteorological Organization, especially in the western TP where the density is less than one-fiftieth of the recommended level[3]. In addition, observations from rain gauges (e.g., siphon, tipping-bucket gauges) may be underestimated, influenced by evaporation, splashing, side wetting, and particularly wind[3]. State-of-the-art satellite precipitation retrievals, such as Global Precipitation Measurements (GPM) Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (IMERG), can provide full coverage over the entire TP. However, previous studies identify notable biases on a daily scale over the TP[10–13]. Reanalysis-based product, such as the fifth-generation ECMWF atmospheric reanalysis (ERA5), offers consistent and continuous

[1]School of Geography and Ocean Science, Nanjing University, Nanjing, China. [2]College of Civil Engineering, Yancheng Institute of Technology, Yancheng, China. [3]Frontiers Science Centre for Critical Earth Material Cycling, Nanjing University, Nanjing, China. [4]Northwest Engineering Corporation Limited, Xi'an, China. [5]Department of Civil Engineering, The University of Hong Kong, Hong Kong SAR, China. [6]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden. [7]Department of Earth System Sciences, Tsinghua University, Beijing, China. [8]Department of Hydraulic Engineering, Tsinghua University, Beijing, China. ✉e-mail: yanglong@nju.edu.cn

Reference	Regional model	Spatial Resolution	Temporal Coverage	Spatial Coverage
Collier *et al.*[28,a]	ERA5-driven COSMO-CLM, ICON-CLM, MPAS, WRF	2.2–4 km	October 2019–September 2020	25°N–40°N, 70°E–115°E
Prein *et al.*[56,b]	ERA5-driven MPAS,RegCM,COSMO-CLM, ICON, WRF	2.2–4 km	July 14–24, 2008; July 27–September 1, 2014; October 1–9 2018	25°N–40°N,70°E–115°E
Ma *et al.*[57]	ERA5-driven WRF	3 km	2009–2018 (June–August)	22°N–43°N, 73°E–104°E
Liu *et al.*[58]	ERA5-driven WRF	4 km	April–September 2017	23°N–44°N, 70°E–108°E
Ma *et al.*[59]	ERA5-driven WRF	3 km	2016–2018 (June–August)	22°N–43°N, 73°E–104°E
Sugimoto *et al.*[60]	ERA5-driven WRF	2 km	2003–2010 (June–September)	28°N–29°N, 86°E–87°E
Cai *et al.*[61]	GFS-driven WRF	4 km	2013–2018 (June–August)	29°N–34°N, 96°E–102°E
Zhou *et al.*[35]	ERA5-driven WRF	0.033° (~4 km)	June–September 2013	25°N–40°N, 70°E–110°E
Gao *et al.*[25], Zhao *et al.*[26]	ERA-Interim driven WRF	4 km	October 2013–May 2014	26°N–40°N, 75°E–105°E
Li *et al.*[62]	GA6.1-driven MetUM	4 km	April–September 2009	17°N–44°N, 70°E–139°E
Lin *et al.*[63]	ERA-Interim driven WRF	2 km	June–August 2015	26°N–28°N, 86°E–90°E
Collier & Immerzeel[29]	ERA-Interim driven WRF-CMB	1 km	17 June 2012–16 June 2013	27°N–29°N, 84°E–86°E

Table 1. Summary of existing convection-permitting simulations over TP. [a]It is an ensemble of thirteen kilometer-scale regional simulations conducted by 10 international research groups. [b]It is an ensemble of seven kilometer-scale regional simulations, as shown in Table 1 of Prein *et al.*[56].

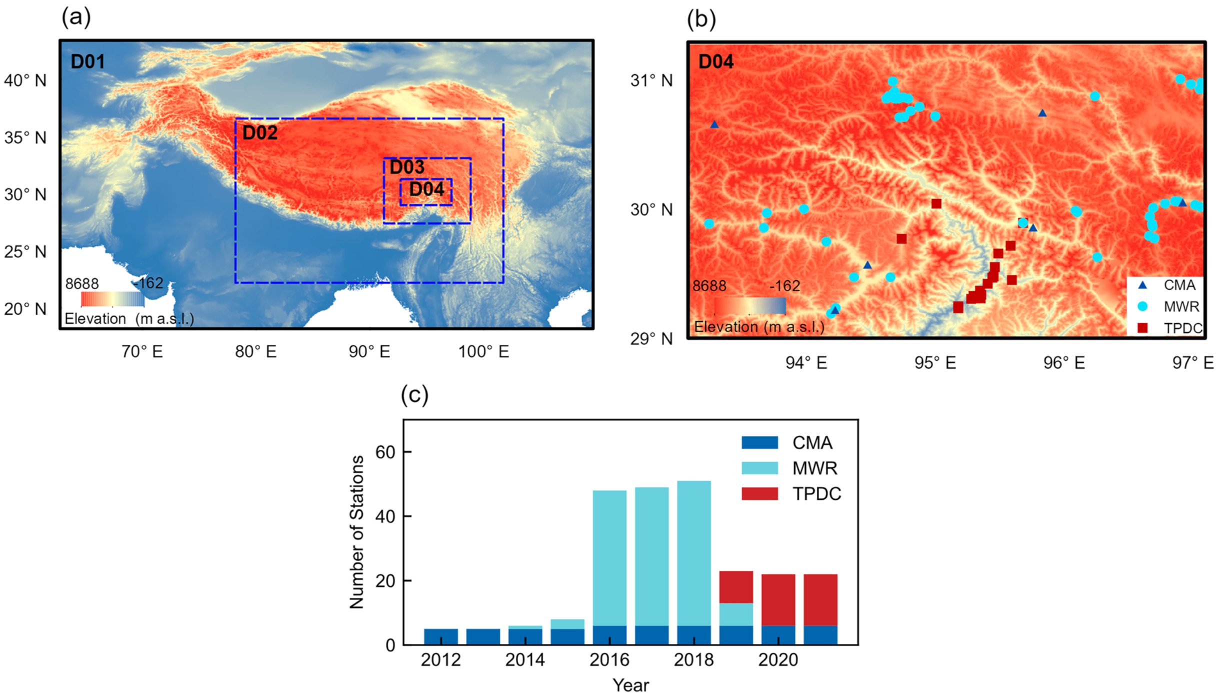

Fig. 1 (**a**) Study area and the four one-way nested domains for the Weather Research and Forecasting model; (**b**) Elevation over domain 4 with 67 rain gauges, and (**c**) annual number of rain gauges from three different sources (i.e., CMA, MWR, TPDC) during 2012–2021.

alternatives. However, the reanalysis products overestimate precipitation amounts exceeding 1 mm/d due to coarse spatial resolution in climate models and their inherent uncertainties of parametrization schemes in representing precipitation processes[14]. Merging multiple datasets demonstrates an effective endeavor to obtain reliable precipitation estimates[10,15–17]. For instance, Li *et al.*[18] merged three satellite precipitation products with a dense network of rain gauge observations based on dynamic Bayesian model averaging. The merged product enhances the accuracy of precipitation accumulation and detection of gauged precipitation events. However, the spatial and temporal resolutions of the existing precipitation products are no less than 0.1° and daily. There are scarce exceptions that the spatial resolutions can be marginally higher based on relatively dense rain gauges[7].

Due to the imperative need for high-resolution precipitation products, kilometer-scale simulations conducted by convection-permitting models (CPMs, 1–4 km resolution), emerge as significant subjects of interest over TP. Precipitation output based on CPMs is often known as dynamic downscaling[19–23]. This is mainly promoted by the Coordinated Regional Climate Downscaling Experiment (CORDEX) flagship pilot studies over the TP[8,21,24]. However, CPM simulations require substantial computational resources, often limiting the

Description	Configuration
Timing	
Simulation period	2012–2021 (June to September)
Time step	30 s (15 s)
Maps and grids	
Map projection	Lambert conformal
Horizontal grid spacing	27 km (149 × 149), 9 km (237 × 171), 3 km (231 × 198), and 1 km (417 × 228)
Vertical levels	38
Forcing strategy	
Initial and boundary conditions	ERA5 (0.25 degree, hourly)
Initialization	Monthly
Runs duration	a calendar month
Spin-up time	3 days before a calendar month
Physical parameterization schemes	
Shortwave radiation	Dudhia scheme
Longwave radiation	Rapid Radiative Transfer Model (RRTM)
Cumulus parameterization	Kain-Fritsch (new Eta) scheme
Microphysics	WSM 3-class simple ice scheme
Land surface model	Noah land-surface model
Planetary boundary layer	YSU scheme
Nesting strategy	One-way nesting
Nudging strategy	Spectral nudging

Table 2. WRF model configuration.

simulation period in many studies to less than one year. For instance, Gao *et al.*[25] conducted model simulations using the Weather Research and Forecasting (WRF) model with a 4 km horizontal grid spacing for the period October 2013 to May 2014 over TP. They show added values, evidenced by lower bias and higher pattern correlation from CPMs when compared against a 28-km WRF simulation and three merged gridded datasets. Zhou *et al.*[26] conducted a high-resolution WRF simulation with a grid spacing of 0.033° (~4 km) from June to September 2013 over TP, with superior accuracy of simulated 10-m wind and precipitation compared to ERA5 and the updated High Asia Refined regional reanalysis (HAR, version 2, hereafter HAR v2)[27]. Collier *et al.*[28] presented an ensemble of thirteen CPM simulations with a horizontal grid spacing ranging from 2.2 to 4 km, covering one hydrological year from October 2019 to September 2020 over TP.

However, existing CPM simulations over the TP mostly have resolutions exceeding 2 km and span a very short period (see Table 1 for a summary). The complex topography over the TP, with narrow valleys often less than 2 km wide, poses significant challenges for accurately capturing local features using existing CPMs' model configurations[29]. A sufficient long span of precipitation product is critical for both hydroclimatological and hydrological analyses over the TP.

In this study, a 1 km-resolution Hourly pRecipitation dataset over the southeastern TP, termed HPTP-1 km, was generated using the WRF model. The dataset spans from 2012 to 2021 for the warm season (June to September). To our knowledge, this is the first time to generate a relatively long-term, fine-resolution sub-daily precipitation product over TP. We focus on southeastern TP as the domain of our pilot study mainly because it is the home to the world's deepest and longest canyon[30]. The strong topographic relief makes the domain the targeted region for hydropower development. The southeastern TP region also experiences the most frequent convective activities during the warm season that act as agents for flash floods and debris flows[31]. This region is also recognized as one of the world's biodiversity hotspots[32]. Better characterization of precipitation variabilities based on a high-resolution precipitation product is important for maintaining water security and sustainable development goals over the TP.

The entire WRF simulation approximately consumes a total of 780,000 CPU core hours. The benefits of convection-permitting simulations and the new precipitation product are highlighted by comparing them against a dense network of *in-situ* precipitation observations, and three state-of-the-art gridded precipitation products, including HAR v2 simulations, ERA5-Land reanalysis, and IMERG satellite retrievals.

Methods

Study area. Our study area mainly covers the southeastern TP as shown in Fig. 1a and b. As we will introduce in section 2.3, the innermost WRF domain (D04) centers over downstream of Yarlung Zangbo River along the section of the Tsangpo Grand Canyon and Motuo County. The Tsangpo Grand Canyon functions as a pathway for water vapor transport, enabling the influx of Asian summer monsoons into the interior of TP[30]. The total area is approximately 0.11 million km^2, including the Parlung Tsangpo River basin, one of the main tributaries for the Yarlung Zangbo River.

Precipitation observations. *In-situ rain gauge observations.* We obtain *in-situ* rain gauge observations from three different sources, i.e., the China Meteorological Administration (CMA, https://data.cma.cn)[18],

Metric	Equation	Unit	Range, perfect
Mean Absolute Error (MAE)	$\text{MAE} = 1/N\sum_{i=1}^{N}\lvert S_i - O_i\rvert$	mm	$[0, +\infty]$, 0
Root Mean Squared Error (RMSE)	$\text{RMSE} = \sqrt{1/N\sum_{i=1}^{N}(S_i - O_i)^2}$	mm	$[0, +\infty]$, 0
Bias	$\text{Bias} = \sum_{i=1}^{N}(S_i - O_i)$	mm	$[-\infty, +\infty]$, 0
Probability of Detection (POD)	$\text{POD} = \frac{Hits}{Hits + Misses}$	—	[0, 1], 1
False Alarm Ratio (FAR)	$\text{FAR} = \frac{False\ Alarms}{Hits + False\ Alarms}$	—	[0, 1], 0
Critical Success Index (CSI)	$\text{CSI} = \frac{Hits}{Hits + Misses + False\ Alarms}$	—	[0, 1], 1

Table 3. Statistical Evaluation metrics for the four precipitation datasets and rain gauge observations. *Note*: N is the number of samples, S_i and O_i are the rainfall simulations and rain gauge observations, respectively; *Hits* are the number of correctly simulated events; *Misses* are the number of events that were observed but were not simulated; *False Alarms* are the number of events that were simulated but did not observe.

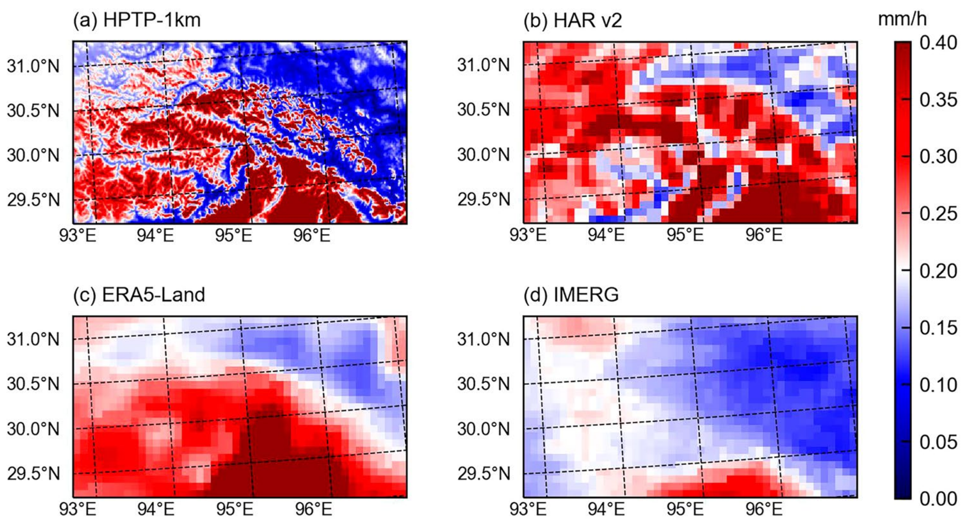

Fig. 2 Spatial distribution of mean hourly precipitation rate from (**a**) HPTP-1km, (**b**) HAR v2, (**c**) ERA5-Land, and (**d**) IMERG during 2012–2021. The spatial resolution of HAR v2, ERA5-Land, and IMERG is approximately 10 km.

the Ministry of Water Resources of China (MWR, http://www.mwr.gov.cn), and the National Tibetan Plateau Data Center (TPDC, http://data.tpdc.ac.cn)[30,33,34]. The number of available rain gauges varies across the period (Fig. 1c). CMA provides a comprehensive precipitation dataset for China, covering the period from 1950 to the present. Most prior studies in the TP have relied on precipitation observations from CMA[22,35], although the rain gauge networks are often considered coarse. In 2014, MWR constructed an extensive network of rain gauges to enhance the monitoring and alert systems for flash floods over TP[18]. Rain gauge data from MWR spanning six years (2014–2019) are obtained over TP. The TPDC provides precipitation observations from November 2019 to December 2022. All *in-situ* rain gauge observations are recorded using the Local Standard Time (LST, UTC + 8 hours). The precipitation observations from CMA and MWR are on a daily scale. A day is defined as from 2000 LST to 2000 LST of the following day. The TPDC data is initially collected hourly and is aggregated to a daily scale.

Gridded precipitation products. We utilize three most popular gridded precipitation products over TP. They are HAR v2, ERA5-Land, and IMERG. HAR v2 is generated using the WRF model and is available at Technische Universität Berlin (https://www.tu.berlin/en/klima/research/ regional-climatology/high-asia/har)[27]. It is an update of the widely used High Asia Refined analysis. The HAR v2 precipitation product has a 10-km resolution, spanning from 1980 to 2023. The ERA5-Land is a global reanalysis dataset generated by re-running the land component of the ERA5 climate reanalysis. It has hourly precipitation data with a spatial resolution of 0.1° and covers the period from 1950 till present. Both ERA5-Land and ERA5 are provided by the European Centre for Medium-Range Weather Forecasts (https://www.ecmwf.int). IMERG is derived from a combination of precipitation-focused satellite passive microwave sensors within the GPM constellation. We utilize the IMERG version 6 Final L3 Half Hourly precipitation product[36] and extract the calibrated precipitation (*precipitationCal*). IMERG provides global half-hourly precipitation data at a resolution of 0.1°, covering the period from June 2000 to September 2021. IMERG precipitation can be accessible through NASA's Earth Science Data collections (https://www.earthdata.nasa.gov). These products are converted to LST for easy manipulation. The gridded products are resampled to a consistent 10-km resolution based on the nearest neighbor interpolation

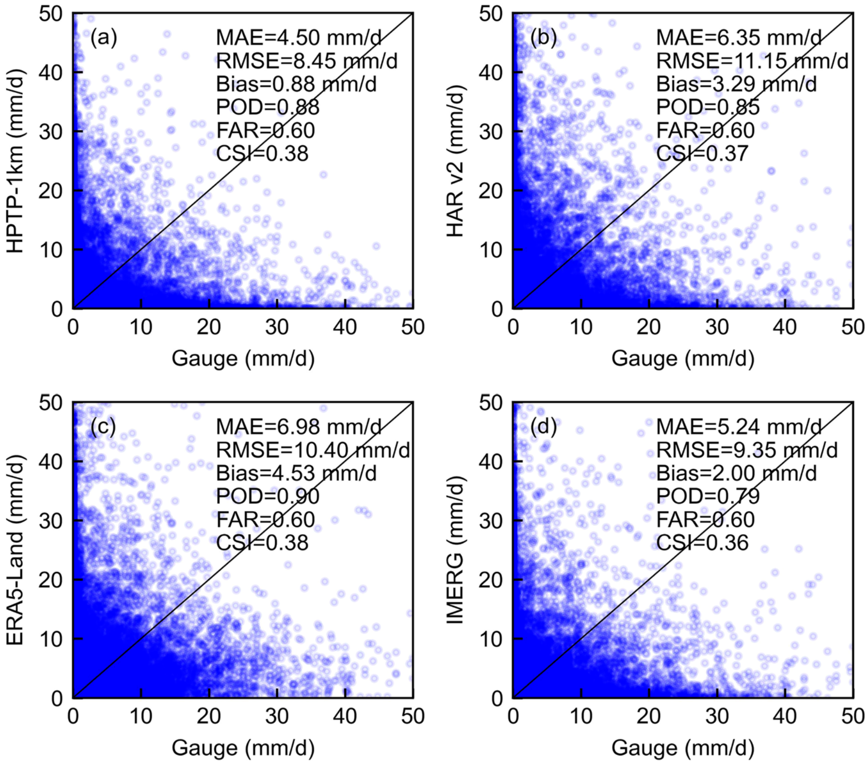

Fig. 3 Daily precipitation from 67 rain gauges compared with the corresponding grids of (**a**) HPTP-1 km, (**b**) HAR2, (**c**) ERA5-Land, and (**d**) IMERG. Six metrics for performance evaluation are shown. The 45° solid line represents the 1:1 reference line.

approach through the Python package xESMF (https://doi.org/10.5281/zenodo.4294774). Our study focuses on the warm seasons with the limited probability of snowfall in this region.

WRF model configuration. Our simulations are based on the Advanced Research configurations of WRF model version 3.9.1. We configured four one-way nested domains (Fig. 1a). The horizontal grid configurations are 149 × 149, 237 × 171, 231 × 198, and 417 × 228 with corresponding grid spacings of 27 km, 9 km, 3 km, and 1 km, respectively. The physics options of the simulations are summarized in Table 2. These options are mainly adapted from previous studies that demonstrate good performance[2,27,37–39]. The initial and boundary conditions were obtained from the ERA5 reanalysis product, with a spatial resolution of 0.25° and hourly temporal resolution.

We conducted a 10-year simulation (2012–2021) and focused on the warm season (i.e., June to September) of each calendar year. In total, this includes simulations spanning 40 months across multiple warm seasons. We implemented a calendar-month reinitialization strategy. Each monthly simulation was initialized at 0000 UTC with a three-day spin-up period preceding the actual calendar month. For example, the 29th, 30th, and 31st of May were designated as the spin-up days for the June simulation. We implemented spectral nudging at 6-hour intervals to enhance long-duration simulations. Spectral nudging is a data assimilation technique to avoid the simulations being "drifted" away from boundary conditions[40,41]. The coefficient of the spectral nudging parameterization is $3.0 \times 10^{-4} s^{-1}$ for temperature, wind components, and moisture. Again, these are recommended by referring to previous studies[40,41]. The time step for all simulations was initially set as 30 seconds. A reduced time step to 15 seconds was implemented when the simulations encountered numerical instability. This is mainly due to complex terrain and/or fast wind speed that call for a short time of integration. We carried out multiple computational tests and settled down with 768 CPU cores as the most efficient. Each month-long simulation can be completed in approximately 25.2 hours. The outputs were generated at hourly intervals. The Python packages Xarray[42] and Salem (https://doi.org/10.5281/zenodo.4635291) were utilized to process and visualize the outputs. Our analyses are based on outputs from the innermost domain (1 km) unless otherwise noted. The total precipitation was calculated using summing the variables RAINC and RAINNC from the WRF outputs.

Data Records

HPTP-1 km[43] spans the warm seasons (June to September) from 2012 to 2021 over the southeastern Tibetan Plateau (TP). This dataset is freely available for download at Figshare[43]: https://doi.org/10.6084/m9.figshare.26211149.v1. It is provided in the NETCDF version 4 format with a total data volume of 10.4 GB.

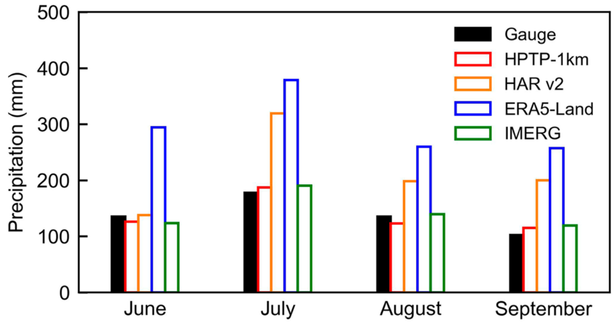

Fig. 4 Ten-year mean monthly precipitation accumulation for the rain gauges and the corresponding grids of HPTP-1 km, HAR v2, ERA5-Land, and IMERG.

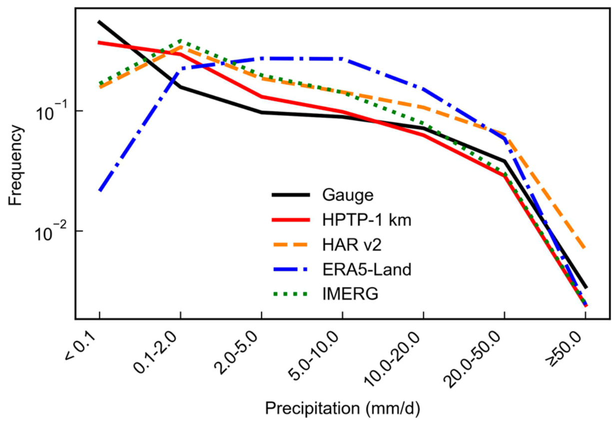

Fig. 5 Histogram of daily precipitation from 67 rain gauges compared with the corresponding HPTP-1km, HAR v2, ERA5-Land, and IMERG.

Technical Validation

We evaluated the performance of gridded precipitation products by comparing them against *in-situ* rain gauge observations using six performance metrics (Table 3). The performance assessment was conducted on a daily scale mainly due to the scarcity of publicly available hourly rain gauge observations. Point values at station locations, extracted from the gridded data using the nearest neighbor interpolation method, were compared with gauge observations. Three accuracy metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Bias, measured the differences between simulations and observations. Three binary classification metrics, including Probability of Detection (POD), False Alarm Ratio (FAR), and Critical Success Index (CSI), assessed the reliability of detecting simulated events calculated from contingency table statistics[44,45]. A commonly used threshold of 1 mm/d was applied to compute POD, FAR, and CSI[44,45]. Details about these metrics are provided in Table 3.

Spatial precipitation patterns. Figure 2 shows the spatial distribution of mean hourly precipitation for the four gridded products across the innermost domain (D04). The HPTP-1 km product captures the fine-scale structure of precipitation distribution, compared to the other three coarse-resolution products (~10 km). The Tsangpo Grand Canyon, as shown in the elevation map (Fig. 1b), comprises a concave ridge with two windward arms that embrace a funnel-shaped valley. HPTP-1 km distinctly delineates the pattern of heavy precipitation on windward slopes and scant precipitation on the lee sides. This aligns with previous research indicating that orographic lifting intensifies extreme precipitation on windward slopes and diminishes it on leeward slopes[46,47]. ERA5-Land and IMERG cannot capture this feature. ERA5-Land and IMERG average precipitation over larger grid cells, smoothing out these localized patterns and failing to represent the sharp gradients of precipitation distribution. HAR v2 compares well with HPTP-1 km. This benefits from the enhanced capability of a regional climate model to capture key processes of orographic precipitation, despite its limitations in resolving finer details (Fig. 2b).

Precipitation comparisons between rain gauges and the corresponding grids. Four gridded precipitation products are compared and evaluated against gauge observations. Note that the gridded precipitation rates at the rain gauge station are determined by calculating the mean within a 3 × 3 window centered on the gauge. This approach mitigates the effects of upwind grids under strong wind conditions over the TP. Figure 3 compares the observed daily precipitation from 67 rain gauges with those over corresponding grids of the four precipitation products. All four gridded precipitation products exhibit very weak correlations with *in-situ* gauge observations. This highlights the challenge of grid-scale characterization of rainfall over TP even for the 1-km

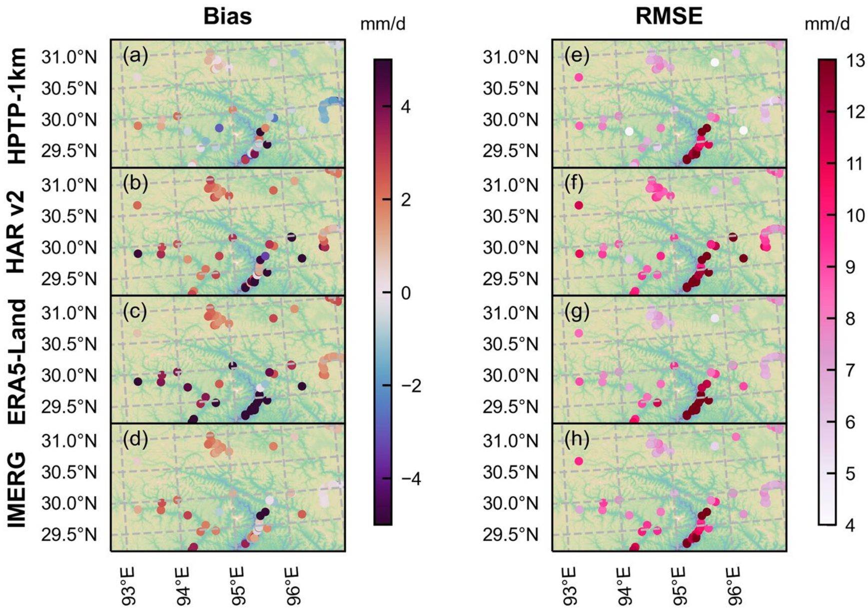

Fig. 6 Spatial distribution of performance metrics for daily precipitation between rain gauges and the corresponding grids from HPTP-1 km, HAR v2, ERA5-Land, and IMERG. (**a**–**d**) Bias, (**e**–**h**) RMSE. The background shade represents elevation over the innermost domain.

convection-permitting model simulations. This echoes Sun *et al.*[21], which shows that a 9-km resolution of daily precipitation from WRF has only an approximate correlation of 0.1 against 37 rain gauges across the entire TP. It suggests that even WRF simulations at a 1-km resolution may still be inadequate for capturing precipitation variability in complex terrain. However, the HPTP-1 km product demonstrates the best performance of all four products, with the smallest values of the three bias metrics (i.e., MAE, RMSE, and Bias).

IMERG and HAR v2 come second and third, respectively, in terms of the bias metrics. Interestingly, the IMERG dataset exhibits significantly lower precipitation rates compared to the other three datasets shown in Fig. 2, yet it demonstrates a higher bias relative to HPTP-1 km. This discrepancy may be attributed to IMERG's underestimation of convective precipitation over mountain ridges, stemming from the limitations of passive microwave sensors (as discussed in Arulraj and Barros[48], Feng *et al.*[49] and Li *et al.*[50]). This results in an overestimation of precipitation rates in valleys and an underestimation over ridges. The ERA5-Land product shows the largest biases. The Bias of ERA5-Land is 4.53 mm/d, exceeding twice that of HPTP-1 km and IMERG. A substantial wet bias in ERA5 has also been identified in previous studies, potentially due to unrealistic lower-level south winds that trigger excessive convective rainfall[21,51]. The new HPTP-1 km product significantly reduces the wet bias. The mean bias is 0.88 mm/d. Both HPTP-1 km and ERA5-Land show comparable binary classification metrics and are slightly better than either IMERG or HAR2.

The superior performance of HPTP-1 km can be further highlighted by comparing the monthly mean precipitation accumulation (Fig. 4). The mean monthly precipitation total for June, July, August, and September derived from the rain gauges (and the corresponding grids from HPTP-1 km) is 135 mm (126 mm), 177 mm (187 mm), 135 mm (123 mm), and 102 mm (115 mm), respectively. IMERG shows comparable performance with that of HPTP-1 km in terms of monthly precipitation total biases. This is possibly due to IMERG undergoing rigorous calibration processes using ground-based observations to capture monthly variations.

To further highlight the capabilities of different precipitation products in characterizing the spectrum of daily precipitation, we compare the probability density distributions of daily precipitation rates between the four gridded precipitation products and rain gauge observations (Fig. 5). Except for HPTP-1 km, the other three gridded precipitation products consistently show large underestimations of the occurrences of light precipitation (i.e., daily precipitation rate smaller than 0.1 mm/d). The reverse is true for moderate precipitation (i.e., daily rain rate between 2.0 mm/d and 20.0 mm/d) where the probability of occurrences is overestimated in HAR v2, IMERG, and ERA5-Land. HPTP-1 km shows the best alignment with that derived from rain gauge observations. In terms of the extreme precipitation rate (i.e., daily precipitation exceeding 20.0 mm/d), the overestimations by HAR v2 persist, while both HPTP-1 km and IMERG show slight underestimation. The overall alignment between HPTP-1 km and rain gauge observations in capturing the spectrum of daily precipitation rates demonstrates the superior performance of the new HPTP-1 km precipitation product to existing gridded precipitation products.

Figure 6 shows the spatial distributions of bias metrics (i.e., Bias and RMSE) for daily precipitation. All the gridded precipitation products show large biases over the Tsangpo Grand Canyon region. More specifically, HAR v2 and IMERG show underestimated daily precipitation (i.e., negative biases) over the canyon, while overestimation is observed over the rest region. Consistent with previous Fig. 3 and Fig. 4, ERA5-Land overall shows overestimation across almost the entire domain. The Bias values are notably reduced in HPTP-1 km. Similarly, all products exhibit large RMSE over the canyon, while relatively low RMSE is observed over the rest region. Notably, the RMSE values are slightly reduced in HPTP-1 km.

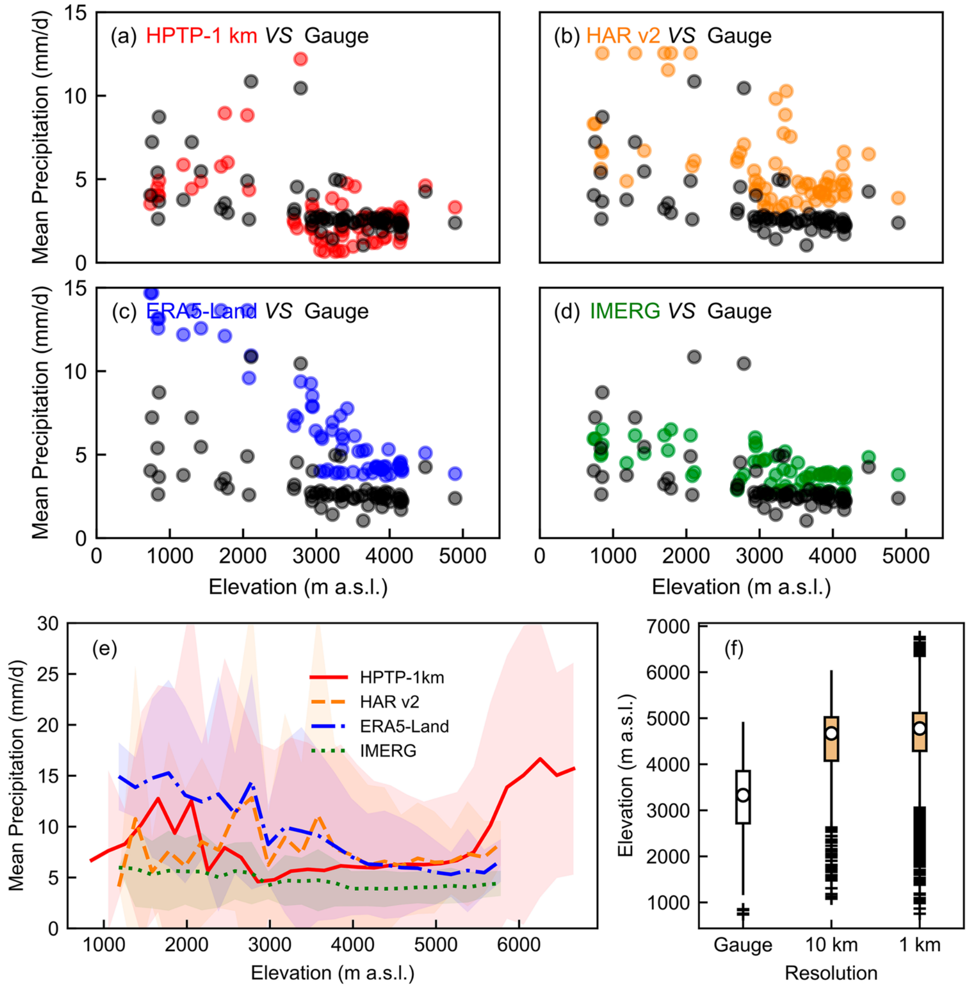

Fig. 7 Elevation dependence of precipitation for gridded data (HPTP-1km, HAR v2, ERA5-Land, and IMERG) and observations over 67 rain gauge stations during 2012–2021. (**a–d**) The mean daily precipitation from gridded data is presented by the values extracted at the rain gauge stations. (**e**) The mean daily precipitation of gridded data is from all grids averaged over 200 m vertical bins during 2012–2021. (**f**) Elevation distributions of rain gauges and grid-based averages across various spatial resolutions. Shaded areas in (**e**) denote the 95% confidence intervals.

Precipitation variation with elevation. Figure 7 compares the mean daily precipitation across different products over the 67 rain gauge stations. Rain gauges over TP are mostly located in low-lying regions[3]. As can be seen from Fig. 7, most of the gauges are located within 3000–4000 m a.s.l., with a small pocket of them located below 2000 m a.s.l. There is an overall tendency of decreased precipitation with elevation, as demonstrated by rain gauge observations. This feature is adequately represented by all four gridded precipitation products (Fig. 7a–d). This is consistent with previous studies and is primarily attributed to lower temperatures, decreased moisture availability, and orographic effects at higher altitudes[12,52]. Specifically, as shown in Fig. 7(a–d), HAR v2 and ERA5-Land show overestimations of precipitation across nearly all elevation ranges. In contrast, HPTP-1 km and IMERG demonstrate better alignment with rain gauge observations, with HPTP-1 km outperforming IMERG, particularly in the elevation range of 3000–4000 m a.s.l. Large variances in precipitation magnitudes are observed below 2000 m a.s.l. This is possibly due to the upward movement of strong, moist air from the South Asian monsoon, which leads to disturbances, cooling, and condensation as the elevation increases. In addition, we compare mean daily precipitation across all grids by categorizing grids into different elevation bands (Fig. 7e). The new HPTP-1 km product shows a pronounced increase in daily precipitation over grids with elevation exceeding 5000 m a.s.l. This is possibly due to the TP vortices presented at the 500-hPa level, with a vertical extent of 2–3 km and a horizontal scale of 400–800 km[53]. These synoptic systems are known as one of the major precipitation-producing systems in the TP. This feature has not been identified based on the other three products, possibly due to their coarse resolutions (Fig. 7f). The non-monotonic dependence of precipitation on elevation increases indicates complexities of spatial precipitation distribution over the southeastern TP.

Diurnal cycle of precipitation. We compare the diurnal cycle as derived from the four gridded precipitation products and *in-situ* rain gauges with hourly precipitation observations (Fig. 8). Hourly values are calculated by averaging precipitation, including zero values, across all days. Hourly observed precipitation peaks at 0200

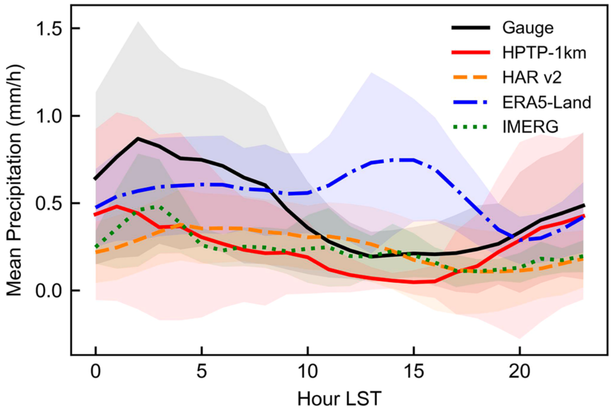

Fig. 8 Diurnal Cycle of precipitation for rain gauges, HPTP-1km, HAR v2, ERA5-Land, and IMERG over 18 rain gauge stations from 2019 to 2021. Shaded areas denote 95% confidence intervals with $\pm 1.96\sigma$. σ denotes the standard deviation of the temporal precipitation series at each hour.

LST over the southwestern TP. The lowest hourly precipitation is observed at around 1300 LST. This finding is consistent with Chen *et al.*[54] which shows precipitation peaks in the early morning and a nadir in the afternoon. The diurnal cycle is well captured by HPTP-1 km, despite the underestimation of hourly precipitation during peak hours. This aligns with Fig. 6 which shows negative biases values. HAR v2 and IMERG capture the diurnal cycle, but demonstrate quite weak amplitudes. For instance, the precipitation peaks and valleys of HAR v2 are 0.38 mm/d and 0.11 mm/d, respectively. ERA5-Land cannot capture the diurnal cycle of precipitation in which the peak time is around 1500 LST compared to the observed early morning peak. This is mainly affected by surface solar heating. However, due to the complexity of topography over the southwestern TP, fine-scale features of topography have a strong modulate of the diurnal cycle of precipitation, such as the mountain-valley breeze. With increased resolutions, there is a significant improvement in HPTP-1 km. However, HPTP-1km still underestimates the peaks of the diurnal cycle, possibly due to existing biases in the initial and boundary fields provided by the ERA5 reanalysis products[30]. The uncertainties in the parameterization schemes of the WRF model could also contribute to the biases, which need additional sensitive experiments in future studies.

Usage Notes

The fine-resolution precipitation product (1 km, hourly), HPTP-1 km, over the southeastern TP for the warm seasons of 2012-2021, is freely accessible at Figshare (https://doi.org/10.6084/m9.figshare.26211149.v1)[43].

Data limitations. We produced a long-term (2012–2021, June to September), fine-resolution (1 km, hourly) precipitation dataset over the southeastern TP (HPTP-1km). However, we acknowledge that the data production and validation have some limitations. Firstly, despite using convection-permitting simulations, some physical processes (e.g., microphysics and turbulence) may still rely on parameterization schemes that introduce uncertainties; variations in nesting and forcing strategies significantly influence model performance[25,55]. Secondly, the accuracy of input data, such as ERA5 reanalysis for initialization and nudging, directly affects the reliability of the simulated precipitation. Errors or biases in the forcing data can propagate into the final product[30]. Thirdly, while 1-km resolution improves representation, it might still fail to capture certain microscale features like sharp ridges, deep gorges, or localized orographic effects[29]. Fourthly, rain gauges are typically located in valleys, which receive less precipitation compared to ridges in the TP. This uneven spatial distribution could bias the validation results and leave some high-altitude regions unverified. Fifthly, the increase in HPTP-1 km precipitation rates at elevations above 5000 m a.s.l. cannot be validated due to a lack of observations at such high altitudes.

Code availability

The HPTP-1km precipitation product was produced using the Weather Research and Forecasting (WRF) model, which is accessible via its Git repository (https://github.com/wrf-model/WRF). Further information can be obtained from the model's official website (https://www2.mmm.ucar.edu/wrf/users). The Python scripts for computation and visualization can be accessible via the associated GitHub repository at https://github.com/liseeocean/WRF-simulations-over-TP.git.

Received: 10 September 2024; Accepted: 14 February 2025;
Published online: 25 February 2025

References

1. Cui, T. *et al*. Non-monotonic changes in Asian Water Towers' streamflow at increasing warming levels. *Nat. Commun.* **14**, 1176 (2023).
2. Yang, L., Ma, J., Wang, X. & Tian, F. Hydroclimatology and hydrometeorology of flooding over the eastern Tibetan Plateau. *J. Geophys. Res.-Atmos.* **127**, e2022JD037097 (2022).
3. Miao, C. *et al*. Understanding the Asian water tower requires a redesigned precipitation observation strategy. *Proc. Natl. Acad. Sci. USA*. **121**, e2403557121 (2024).
4. Yao, T. *et al*. The imbalance of the Asian water tower. *Nat. Rev. Earth Environ.* **3**, 618–632 (2022).

5. Zhou, P., Tang, J., Ma, M., Ji, D. & Shi, J. High resolution Tibetan Plateau regional reanalysis 1961-present. *Sci. Data.* **11**, 444 (2024).
6. Ren, Z. *et al*. Temporal scaling characteristics of sub-daily precipitation in Qinghai-Tibet Plateau. *Earth's Future.* **12**, e2024EF004417 (2024).
7. Jiang, Y. *et al*. TPHiPr: a long-term (1979-2020) high-accuracy precipitation dataset (1∕30°, daily) for the Third Pole region based on high-resolution atmospheric modeling and dense observations. *Earth Syst. Sci. Data.* **15**, 621–638 (2023).
8. Dallan, E. *et al*. How well does a convection-permitting regional climate model represent the reverse orographic effect of extreme hourly precipitation? *Hydrol. Earth Syst. Sci.* **27**, 1133–1149 (2023).
9. Wang, X., Pang, G. & Yang, M. Precipitation over the Tibetan Plateau during recent decades: a review based on observations and simulations. *Int. J. Climatol.* **38**, 1116–1131 (2018).
10. Lyu, Y. & Yong, B. A novel double machine learning strategy for producing high-precision multi-source merging precipitation estimates over the Tibetan Plateau. *Water Resour. Res.* **60**, e2023WR035643 (2024).
11. Lei, H., Li, H., Zhao, H., Ao, T. & Li, X. Comprehensive evaluation of satellite and reanalysis precipitation products over the eastern Tibetan plateau characterized by a high diversity of topographies. *Atmos. Res.* **259**, 105661 (2021).
12. Na, Y., Lu, R., Fu, Q. & Kodama, C. Precipitation characteristics and future changes over the southern slope of Tibetan Plateau simulated by a high-resolution global nonhydrostatic model. *J. Geophys. Res.-Atmos.* **126**, e2020JD033630 (2021).
13. Zhan, C. *et al*. First evaluation of GPM-Era satellite precipitation products with new observations on the western Tibetan Plateau. *Atmos. Res.* **283**, 106559 (2023).
14. Jiang, J. *et al*. Precipitation regime changes in High Mountain Asia driven by cleaner air. *Nature.* **623**, 544–549 (2023).
15. Zhang, L. *et al*. Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach. *J. Hydrol.* **594**, 125969 (2021).
16. Wei, L. *et al*. An Extended Triple Collocation Method With Maximized Correlation for Near Global-Land Precipitation Fusion. *Geophys. Res. Lett.* **50**, e2023GL105120 (2023).
17. Fan, Z. *et al*. A comparative study of four merging approaches for regional precipitation estimation. *IEEE Access* **9**, 33625–33637 (2021).
18. Li, K. *et al*. A high-accuracy rainfall dataset by merging multiple satellites and dense gauges over the southern Tibetan Plateau for 2014–2019 warm seasons. *Earth Syst. Sci. Data.* **13**, 5455–5467 (2021).
19. Politi, N., Vlachogiannis, D., Sfetsos, A. & Nastos, P. T. High-resolution dynamical downscaling of ERA-Interim temperature and precipitation using WRF model for Greece. *Clim. Dyn.* **57**, 799–825 (2021).
20. Jing, X., Geerts, B., Wang, Y. & Liu, C. Evaluating seasonal orographic precipitation in the interior western United States using gauge data, gridded precipitation estimates, and a regional climate simulation. *J. Hydrometeorol.* **18**, 2541–2558 (2017).
21. Sun, H. *et al*. Hydrological evaluation of high-resolution precipitation estimates from the WRF model in the Third Pole river basins. *J. Hydrometeorol.* **22**, 2055–2071 (2021).
22. Ou, T. *et al*. Simulation of summer precipitation diurnal cycles over the Tibetan Plateau at the gray-zone grid spacing for cumulus parameterization. *Clim. Dyn.* **54**, 3525–3539 (2020).
23. Gao, Y., Xu, J. & Chen, D. Evaluation of WRF Mesoscale Climate Simulations over the Tibetan Plateau during 1979–2011. *J. Clim.* **28**, 2823–2841 (2015).
24. Lucas Picher P. *et al*. Convection-permitting modeling with regional climate models: Latest developments and next steps. *WIREs Climate Change*. **12** (2021).
25. Gao, Y., Chen, F. & Jiang, Y. Evaluation of a convection-permitting modeling of precipitation over the Tibetan Plateau and its influences on the simulation of snow-cover fraction. *J. Hydrometeorol.* **21**, 1531–1548 (2020).
26. Zhao, Y., Zhou, T., Li, P., Furtado, K. & Zou, L. Added value of a convection permitting model in simulating atmospheric water cycle over the Asian water tower. *J. Geophys. Res.-Atmos.* **126**, e2021JD034788 (2021).
27. Wang, X., Tolksdorf, V., Otto, M. & Scherer, D. WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis. *Int. J. Climatol.* **41**, 743–762 (2021).
28. Collier E. *et al*. The first ensemble of kilometer-scale simulations of a hydrological year over the third pole. *Clim. Dyn.*, (2024).
29. Collier, E. & Immerzeel, W. W. High-resolution modeling of atmospheric dynamics in the Nepalese Himalaya. *J. Geophys. Res.-Atmos.* **120**, 9882–9896 (2015).
30. Chen, X. *et al*. Investigation of precipitation process in the water vapor channel of the Yarlung Zsangbo Grand Canyon. *Bull. Amer. Meteorol. Soc.* **105**, E370–E386 (2024).
31. Kukulies, J., Chen, D. & Wang, M. Temporal and spatial variations of convection, clouds and precipitation over the Tibetan Plateau from recent satellite observations. Part II: Precipitation climatology derived from global precipitation measurement mission. *Int. J. Climatol.* **40**, 4858–4875 (2020).
32. Wu, P. *et al*. Vegetation classification and distribution patterns in the south slope of Yarlung Zangbo Grand Canyon National Nature Reserve, Eastern Himalayas. *Plants.* **11**, 1194 (2022).
33. Ma Y., Chen X. & Xu X. Investigation the water vapor channel of Yarlung Zangbo Grand Canyon (2018-2019). In: National T. P. D. C. (2024).
34. Chen X. *et al*. Research progress on the water vapor channel within the Yarlung Zsangbo Grand Canyon, China. *Atmos. Ocean. Sci. Lett.*, 100462 (2024).
35. Zhou, X. *et al*. Added value of kilometer-scale modeling over the third pole region: a CORDEX-CPTP pilot study. *Clim. Dyn.* **57**, 1673–1687 (2021).
36. Huffman, G. J. *et al*. Integrated multi-satellite retrievals for the global precipitation measurement (GPM) mission (IMERG). *Satellite precipitation measurement* **ume 1**, 343–353 (2020).
37. Zhang J., Yang L., Yu M. & Chen X. Response of extreme rainfall to atmospheric warming and wetting: implications for hydrologic designs under a changing climate. *J. Geophys. Res.-Atmos*. **128** (2023).
38. Yang, L., Li, K., Shen, Y. & Tian, F. Flood seasonality over the Third Pole Region modulated by upper level moisture transport. *Earth's Future* **10**, e2022EF002828 (2022).
39. Maussion, F. *et al*. Precipitation seasonality and variability over the Tibetan Plateau as resolved by the High Asia Reanalysis. *J. Clim.* **27**, 1910–1927 (2014).
40. Spero, T. L., Nolte, C. G., Mallard, M. S. & Bowden, J. H. A maieutic exploration of nudging strategies for regional climate applications using the WRF model. *J. Appl. Meteorol. Climatol.* **57**, 1883–1906 (2018).
41. Bullock, O. R. *et al*. An observation-based investigation of nudging in WRF for downscaling surface climate information to 12-km grid spacing. *J. Appl. Meteorol. Climatol.* **53**, 20–33 (2014).
42. Hoyer S. & Hamman J. xarray: ND labeled arrays and datasets in Python. *Journal of Open Research Software*. **5** (2017).
43. Li D. A decade-long (2012–2021) fine-resolution (1 km,hourly) rainfall dataset over southeastern Tibetan Plateau. https://doi.org/10.6084/m9.figshare.26211149.v1 (2024).
44. Yong, B. *et al*. Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resour. Res.* **46**, W07542 (2010).
45. Ebert, E. E., Janowiak, J. E. & Kidd, C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteorol. Soc.* **88**, 47–64 (2007).
46. Marra, F., Armon, M., Borga, M. & Morin, E. Orographic Effect on Extreme Precipitation Statistics Peaks at Hourly Time Scales. *Geophys. Res. Lett.* **48**, e2020GL091498 (2021).

47. Singh, P. & Kumar, N. Effect of orography on precipitation in the western Himalayan region. *J. Hydrol.* **199**, 183–206 (1997).
48. Rojas, Y., Minder, J. R., Campbell, L. S., Massmann, A. & Garreaud, R. Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central Chile. *Atmos. Res.* **253**, 105454 (2021).
49. Feng Y., Qi Y., Zhao Z. & Li D. Can Satellite or Reanalysis Precipitation Products Depict the Location and Intensity of Rainfall at Flash Flood Scale Over the Eastern Mountainous Area of the Tibetan Plateau? *Water Resour. Res.* **60** (2024).
50. Li, L. *et al*. Implications for Validation of IMERG Satellite Precipitation in a Complex Mountainous Region. *Remote Sens* **15**, 4380 (2023).
51. Ou, T. *et al*. Wet bias of summer precipitation in the northwestern Tibetan Plateau in ERA5 is linked to overestimated lower-level southerly wind over the plateau. *Clim. Dyn.* **61**, 2139–2153 (2023).
52. Zhang, H. & Gao, Y. Projected changes of precipitation over the Qinghai-Tibetan Plateau based on dynamical downscaling. *Plateau Meteorol* **39**, 477–485 (2020).
53. Curio, J., Schiemannm, R., Hodges, K. I. & Turner, A. G. Climatology of Tibetan Plateau Vortices in Reanalysis Data and a High-Resolution Global Climate Model. *J. Clim.* **32**, 1933–1950 (2019).
54. Chen, X., Cao, D., Liu, Y., Xu, X. & Ma, Y. An observational view of rainfall characteristics and evaluation of ERA5 diurnal cycle in the Yarlung Tsangbo Grand Canyon, China. *Q. J. R. Meteorol. Soc.* **149**, 1459–1472 (2023).
55. Maussion, F. *et al*. WRF simulation of a precipitation event over the Tibetan Plateau, China—an assessment using remote sensing and ground observations. *Hydrol. Earth Syst. Sci.* **15**, 1795–1817 (2011).
56. Prein, A. F. *et al*. Towards ensemble-based kilometer-scale climate simulations over the Third Pole Region. *Clim. Dyn.* **60**, 4055–4081 (2023).
57. Ma, M. *et al*. Summer regional climate simulations over Tibetan Plateau: from gray zone to convection permitting scale. *Clim. Dyn.* **60**, 301–322 (2023).
58. Liu, H., Liu, X., Liu, C. & Yun, Y. High-resolution regional climate modeling of warm-season precipitation over the Tibetan Plateau: Impact of grid spacing and convective parameterization. *Atmos. Res.* **281**, 106498 (2023).
59. Ma, M., Hui, P., Liu, D., Zhou, P. & Tang, J. Convection-permitting regional climate simulations over Tibetan Plateau: re-initialization versus spectral nudging. *Clim. Dyn.* **58**, 1719–1735 (2022).
60. Sugimoto, S. *et al*. Cloud-resolving-model simulations of nocturnal precipitation over the Himalayan slopes and foothills. *J. Hydrometeorol.* **22**, 3171–3188 (2021).
61. Cai, S. *et al*. Diurnal cycle of summer precipitation over the Eastern Tibetan Plateau and surrounding regions simulated in a convection-permitting model. *Clim. Dyn.* **57**, 611–632 (2021).
62. Li, D. *et al*. Characterizing precipitation in high altitudes of the western Tibetan plateau with a focus on major glacier areas. *Int. J. Climatol.* **40**, 5114–5127 (2020).
63. Lin, C., Chen, D., Yang, K. & Ou, T. Impact of model resolution on simulating the water vapor transport through the central Himalayas: implication for models' wet bias over the Tibetan Plateau. *Clim. Dyn.* **51**, 3195–3207 (2018).

Acknowledgements
This study was supported by the National Natural Science Foundation of China (42401045, 52379012), the China Postdoctoral Science Foundation (2023M741637), the Fundamental Research Funds for the Central Universities (0209-14380126, 0209-14380133), the Fundamental Research Funds for the Central Universities-Cemac "GeoX" Interdisciplinary Program (2024QNXZ03), and the funding for school-level research projects of Yancheng Institute of Technology (xjr2023017). FT acknowledges support from the National Key Research and Development Program of China (2022YFC3002900). The numerical simulations in this paper are implemented on the computing facilities in the High-Performance Computing Center (HPCC) of Nanjing University.

Author contributions
Dayang Li: Conceptualization, Methodology, Visualization, Writing- Original draft preparation. Long Yang: Conceptualization, Methodology, Validation, Writing - review & editing, Supervision. Ye Shen, Xianyi Ye & Xiaohua Fu: Methodology, Data curation, Writing - review & editing. Yang Yang, TingHai Ou, Deliang Chen & Fuqiang Tian: Methodology, Writing - review & editing.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to L.Y.

Reprints and permissions information is available at www.nature.com/reprints.

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