Earth's Future

RESEARCH ARTICLE
10.1029/2024EF004417

Temporal Scaling Characteristics of Sub-Daily Precipitation in Qinghai-Tibet Plateau

Zhihui Ren[1,2], Yan-Fang Sang[1,2,3,4], Peng Cui[5], Deliang Chen[6], Yichi Zhang[1,2,3], Tongliang Gong[4,7], Shao Sun[8], and Nedra Mellouli[9]

[1]Key Laboratory of Water Cycle & Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Key Laboratory of Compound and Chained Natural Hazards, Ministry of Emergency Management of China, Beijing, China, [4]Yarlung Zangbo Grand Canyon Water Cycle Monitoring and Research Station, Linzhi, China, [5]Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, [6]Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden, [7]Water Conservancy Project & Civil Engineering College, Tibet Agriculture & Animal Husbandry University, Linzhi, China, [8]State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China, [9]Artificial Intelligence and Data Semantics, Paris 8 University, ESILV-Devinci Group, Paris, France

Abstract The Qinghai-Tibet Plateau (QTP) is highly susceptible to destructive rainstorm hazards and related natural disasters. However, the lack of sub-daily precipitation observations in this region has hindered our understanding of rainstorm-related hazards and their societal impacts. To address this data gap, a new approach is devised to estimate sub-daily precipitation in QTP using daily precipitation data and geographical information. The approach involves establishing a statistical relationship between daily and sub-daily precipitation based on data from 102 observation sites. This process results in a set of functions with six associated parameters. These parameters are then modeled using local geographical and climatic information through a machine learning algorithm called support vector regression. The results indicated that the temporal scaling characteristics of sub-daily precipitation can be accurately described using a logarithmic function. The uncertainty of the estimates is quantified using the coefficient of variance and coefficient of skewness, which are estimated using a logarithmic and linear curve, respectively. Additionally, the six parameters are found to be closely linked to geographical conditions, enabling the creation of a 1-km parameters data set. This data set can be utilized to quantitatively describe the probabilistic distribution and extract key information about maximum precipitation duration (from 1 to 12 hr). Overall, the findings suggest that the generated parameters data set holds significant potential for various applications, including risk analysis, forecasting, and early warning for rainstorm-related natural disasters in QTP. The innovative method developed in this study proves to be an effective approach for estimating sub-daily precipitation and assessing its uncertainty in ungauged regions.

Plain Language Summary As one of famous hotspots for natural disaster studies on Earth, the Qinghai-Tibet Plateau (QTP) is highly vulnerable to destructive rainstorm hazard and related natural disasters, causing significant damage to property, infrastructure, agriculture, and resulting in extensive loss of life. Short-duration heavy precipitation at sub-daily scales is an important trigger for flash flood, debris flows and other disasters in QTP. However, it is a poorly gauged high mountain region, observed data for sub-daily precipitation is extremely limited. Although there have been several satellite products and reanalysis data for sub-daily precipitation in QTP, their quality has large bias and uncertainty compared to observations. It leaves a large data gap of sub-daily precipitation, hindering the studies of rainstorm-related natural disasters in the region. In this work, we develop a new strategy to quantify the temporal scaling characteristics of sub-daily precipitation, as a basis of temporal downscaling. Then we use the new strategy to generate a parameters data set, to fill the data gap of sub-daily precipitation in QTP. The parameters data set generated provides an effective way to estimate sub-daily precipitation and its uncertainty, which can effectively serve for the rainstorm-related natural disasters study in QTP.

Key Points:

- A high-resolution gridded parameters data set is generated to estimate sub-daily precipitation and its uncertainty in QTP
- The temporal scaling characteristics of sub-daily precipitation in QTP is well described by a logarithmic function
- Spatial heterogeneity in the temporal scaling characteristics of sub-daily precipitation in QTP is closely related to geographical conditions

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y.-F. Sang,
sangyf@igsnrr.ac.cn;
sunsangyf@gmail.com

Citation:

Ren, Z., Sang, Y.-F., Cui, P., Chen, D., Zhang, Y., Gong, T., et al. (2024). Temporal scaling characteristics of sub-daily precipitation in Qinghai-Tibet Plateau. *Earth's Future*, *12*, e2024EF004417. https://doi.org/10.1029/2024EF004417

Received 13 JAN 2024
Accepted 11 FEB 2024

Author Contributions:

Conceptualization: Zhihui Ren, Yan-Fang Sang
Data curation: Yan-Fang Sang, Yichi Zhang, Shao Sun
Formal analysis: Zhihui Ren
Funding acquisition: Yan-Fang Sang
Investigation: Peng Cui, Deliang Chen, Yichi Zhang, Tongliang Gong, Shao Sun, Nedra Mellouli
Methodology: Zhihui Ren, Yan-Fang Sang, Peng Cui, Yichi Zhang, Nedra Mellouli
Project administration: Yan-Fang Sang
Resources: Yan-Fang Sang
Software: Zhihui Ren, Yan-Fang Sang

1. Introduction

High mountain regions (HMR) are highly vulnerable to destructive rainstorm hazard and related natural disasters, including floods, landslides, erosion, mud and debris slides (Hock et al., 2022; Yao et al., 2013). These hazards

Supervision: Yan-Fang Sang, Peng Cui, Deliang Chen
Validation: Zhihui Ren, Yan-Fang Sang, Peng Cui, Deliang Chen
Visualization: Zhihui Ren, Yan-Fang Sang
Writing – original draft: Zhihui Ren
Writing – review & editing: Yan-Fang Sang, Peng Cui, Deliang Chen, Yichi Zhang, Tongliang Gong, Shao Sun, Nedra Mellouli

pose a significant threat to human lives, properties, the environment, and ecosystems in both HMR and downstream regions (Terzi et al., 2019; Viviroli et al., 2007). Moreover, due to the high elevations, the temperature in HMR usually has a higher increasing rate compared to plain regions (Pepin et al., 2015, 2022), thus HMR are more sensitive to the impacts of climate change, due to which the risks of rainstorm-related natural disasters keep increasing and their potential influences are aggravating (Immerzeel et al., 2019). However, effective control and prevention of rainstorm-related natural disasters in HMR is a big challenging issue. On the one hand, it is due to the unfavorable characteristics of these disasters, such as sudden occurrence along with high-intensive rainfall, rapid formation and evolution of disasters, and strong destructive outcomes, leading to high difficulty in simulation and forecasting of natural disasters in HMR. On the other hand, due to the high elevation and complex topography and harsh geographical conditions, field hydro-climatic observations, especially short-duration rainfall observations, are barely available in HMR, hindering the studies of natural disasters. Hence, how to accurately evaluate the risks of rainstorm-related natural disasters in HMR, as well as their reliable forecasting and early warning, is a high requirement but also an extraordinarily challenging task.

Qinghai-Tibet Plateau (QTP) is a typical HMR with damaging natural disasters occurring frequently. In the region, short-duration heavy precipitation at sub-daily scales is an important trigger not only for flash flooding disasters but also for landslide and debris flows (Zhao et al., 2022). Over the last two decades, more than 2,000 rainstorm-induced disasters occurred across the region, and they affected 7.2 million people and caused economic losses of 13.9 billion RMB (Sun et al., 2021). However, QTP is also a typical ungauged region, and short-duration precipitation data with long observation periods are limited, leaving a large data gap for supporting rainstorm-related natural disaster studies in the region.

Satellite remote sensing data and reanalysis data are important alternative sources for obtaining precipitation data in ungauged regions. The two data sources provide good spatial coverage, however, their quality relies heavily on the inversion techniques and requires bias correction with the help of rain gauges (Lu & Yong, 2018). Several daily precipitation data sets are available for QTP, and previous studies made deep comparisons and improvements in their quality (Y. Ma et al., 2018; Tong et al., 2014). These daily data, however, are too coarse to meet the high time-resolution request at sub-daily scales for the rainstorm-related natural disaster studies in the region (Jiang et al., 2023; Wu & Gao, 2013). Although there have been several satellite products for sub-daily precipitation in QTP, their quality is much lower than that of daily precipitation products (Z. Ma et al., 2022; Sun et al., 2018; G. Tang et al., 2016), due to the lack of field sub-daily precipitation observations for calibration, and thus they cannot reliably support the rainstorm-related natural disaster studies (Casanueva et al., 2020).

Temporal downscaling is another way to generate sub-daily precipitation data from daily precipitation, and dynamical downscaling and statistical downscaling are two approaches used widely (J. Tang et al., 2016; Yan et al., 2021; Yin & Chen, 2020). Dynamical downscaling is computationally expensive because of a large amount of computation and is greatly influenced by the boundary conditions provided by global climate models or regional climate models (Maraun et al., 2010; Ou et al., 2020). Comparatively, statistical downscaling is based on the establishment of statistical relationships between local observations and large-scale climatic (sometimes also geographical) variables, and is computationally effective (Benestad et al., 2008; Chen et al., 2005; Hanssen-Bauer et al., 2005). Thus, this approach is an effective alternative for generating high-time-resolution precipitation data in ungauged regions (Wilby et al., 1998).

In the framework of statistical downscaling, many machine learning and deep learning methods were used to explore the connections between low- and high-time resolution scales, such as the artificial neural network method (Vu et al., 2015), fractal-multifractal method (Maskey et al., 2019), multiplicative cascade method (Müller & Haberlandt, 2018), extreme learning machine method (MoradiKhaneghahi et al., 2019), and stochastic disaggregation method (Socolofsky et al., 2001; Takhellambam et al., 2022). The above temporal downscaling methods usually pursue a statistically perfect fit; however, their effectiveness depends not only on the data availability but also on physical mechanisms of precipitation. Some studies used statistical methods for temporal downscaling based on probabilistic distribution theory, such as exponential distribution (Benestad et al., 2012, 2021)and generalized extreme value distribution (Lima et al., 2018). The relative consistency in precipitation frequencies at different timescales is employed to make a temporal downscaling, but it is limited to the context at a specific frequency. Besides, other studies also used the linear scaling methods for temporal downscaling from daily to hourly scales (Cattoën et al., 2020; Lee & Jeong, 2014; Lee & Park, 2017), but ignoring the underlying nonlinear relationship.

The quality of any empirical temporal downscaling strongly depends on the number of field observations and the statistical methods used. Compared to the complex and indescribable relationship between daily and sub-daily precipitation generated by machine learning and deep learning methods in previous studies (Müller & Haberlandt, 2018; Socolofsky et al., 2001; Takhellambam et al., 2022), it could be a more effective and feasible strategy to establish a describable equation of the relationship of precipitation between daily and sub-daily scales, to quantify the temporal scaling characteristics of sub-daily precipitation. The motivation for this strategy is the observation and expectation that the temporal scaling characteristics of sub-daily precipitation has simple and smooth shapes with regional similarities (Parding et al., 2023). Based on the describable equation established, it would be favorable to understand the spatial pattern of sub-daily precipitation and its influencing factors in a region.

The main objective of this study is to generate a high-resolution gridded parameters data set to quantify the temporal scaling characteristics of sub-daily precipitation and the associated uncertainty in the Qinghai-Tibet Plateau, as a basis of estimating sub-daily precipitation in the region. To do this, the hourly precipitation data measured at 102 weather stations in QTP are used, and a logarithmic function, with two key parameters, is selected to describe the temporal scaling characteristics of sub-daily precipitation at these stations. Furthermore, the influences of the geographical and climatic conditions on the two key parameters at these stations are further explored, based on which a 1-km gridded data set of the two key parameters is generated in QTP. To the best of our knowledge, exploring the temporal scaling characteristics of sub-daily precipitation in QTP is a novel approach. The parameters data set provides an effective and practical approach to estimating sub-daily precipitation and the associated uncertainty for areas with available daily precipitation data, supporting the rainstorm-related natural disaster studies in QTP.

2. Study Area and Data

2.1. Study Area

The Qinghai-Tibet Plateau, well known as "the Third Pole," is located in the south-central part of the Eurasian continent, between 25°59′N–40°1′N and 67°40′E–104°40′E and covering an area of over 3 million km^2 (Y. Zhang et al., 2021). It is the highest (average altitude >4,000 m asl) and largest plateau in the world (G. Zhang et al., 2013). The QTP is characterized by plains, valleys, and mountains, and the surface topography changes from very gentle to highly steeped conditions (0–78°) (Sajadi et al., 2022). Under the influences of regional monsoon systems, precipitation from May to September accounts for more than 70% of annual precipitation in most parts of QTP (Kukulies et al., 2020; Zhu et al., 2020).

The east and south parts of QTP, with extremely valley topography, are the origins of several large Asian rivers including the Yellow River, Yangtze River, Mekong River, Salween River Basin, Brahmaputra River and others (Liu et al., 2020). The climatic conditions in these river valleys are controlled by the regional monsoon systems, and thus rainstorms frequently occur in summer, generating serious flash floods, landslides, mudslides and other natural disasters, as one of the hotspots for natural disaster studies on Earth. Unfortunately, the population and economy are also concentrated in these river valleys and thus are prone to destructive rainstorm-related natural disasters (Cui & Jia, 2015; W. Ma et al., 2021). Under the influences of climate change, the occurrence frequency and intensity of natural hazards keep increasing, which poses serious threats to people's lives and properties in QTP.

In this study, we mainly concern the east and south parts of QTP, and it includes eight basins: the Hexi Corridor Basin, the Qaidam Basin and the origin of the Yellow River Basin in the northeast, the origin of the Jinsha River Basin, the Mekong River Basin and the Salween River Basin in the southeast, the Yarlung Tsangpo River Basin (as the upper reach of the Brahmaputra River Basin), and the origin of the Ganges River Basin in the south (see Figure 1).

2.2. Data

2.2.1. Hourly Precipitation Data

The hourly precipitation data measured at 102 weather stations (see Figure 1) in QTP are used for this study. They were collected from the China Meteorological Administration (CMA). All the data have been quality checked to ensure their reliability for scientific studies. The weather stations are mainly located in the eight basins concerned

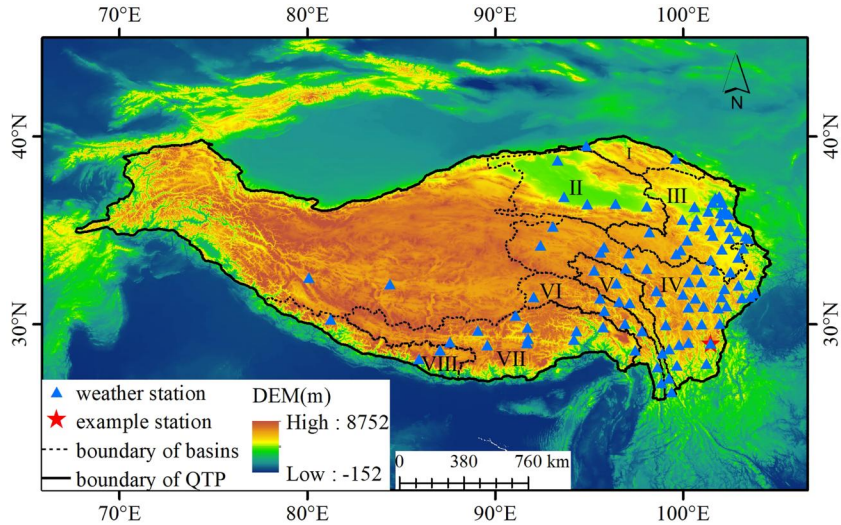

Figure 1. Locations of the Qinghai-Tibet Plateau (QTP) and the 102 weather stations used in this study. The station with the star symbol is used as an example for the analysis in Section 4.1. I: the Hexi Corridor Basin; II: the Qaidam Basin; III: the origin of the Yellow River; IV: the origin of the Jinsha River Basin; V: the origin of the Mekong River Basin; VI: the origin of the Salween River Basin; VII: the Yarlung Tsangpo River Basin; VII: the origin of the Ganges River Basin.

in this study, where rainstorm-related natural disasters occur frequently (Wang et al., 2021; Yang et al., 2022). However, there are limited stations in the other regions of QTP, where historical occurrences of rainstorm-related natural disasters were much less frequent than in the river valley regions. These hourly precipitation data have different measured periods, with different starting years from 1951 to 2012, but with the same ending year of 2018. For each station, all the measured data are used to investigate the relationship between daily and sub-daily precipitation. Considering that rainstorm hazard and related natural disasters mainly occur in summer in QTP, the hourly precipitation data from May to September in each year are used for this study.

Based on the Universal Time Coordinated (UTC) time, the daily precipitation (denoted as P_{24}) is calculated as the sum of hourly precipitation from 0:00 UTC to 24:00 UTC. Only wet days with cumulative precipitation exceeding 1.0 mm are counted here. For certain station, the maximum magnitude of t-hour consecutive precipitation (denoted as P_t, with $t = 1, 2, \ldots, 12$ hr, respectively) in a wet day is calculated to establish its $P_t/P_{24} \sim t$ curve; after that, the average curve of these $P_t/P_{24} \sim t$ curves among all wet days is calculated, to quantify the temporal scaling characteristics of sub-daily precipitation at the station.

Moreover, the uncertainty of the average $P_t/P_{24} \sim t$ curve is quantified by using the indicators of coefficient of variance (C_v), and coefficient of skewness (C_s), which can also be estimated from all the $P_t/P_{24} \sim t$ curves at a station. The different $P_t/P_{24} \sim t$ curves at the 102 weather stations are further compared to investigate the spatial difference.

The ratios of sub-daily precipitation ($t = 1, 3, 6, 12$ hr) to daily precipitation at these stations are used as examples to present the regional similarities. Figure 2 clearly indicates that the ratio of sub-daily precipitation to daily precipitation is larger in the northern part than in the southern part of the study area. For example, P_1 accounts for less than 46% (Figure 2a) of P_{24} in the southeastern part of QTP, but more than 46% in other regions. The spatial distribution of P_3/P_{24} ratio (Figure 2b) is similar to that of P_1/P_{24}. Besides, the P_6/P_{24} (Figure 2c) and P_{12}/P_{24} (Figure 2d) ratios are larger in the northern part than in the southern part of the study area, implying that heavy precipitation generally has a longer time duration in the southeastern QTP with wet climatic conditions. Overall, the duration of precipitation events is usually no longer than 12 hr in the study area, as the P_{12}/P_{24} ratios are larger than 86% at all stations. Note that it is the reason of setting the longest $t = 12$ hr in this study.

The regional similarities of P_t/P_{24} shown in Figure 2 imply the potential influence of geographical and climatic conditions on the sub-daily precipitation ratios. This motivates investigating the temporal scaling characteristics of sub-daily precipitation, as a new and feasible strategy developed in this study to estimate sub-daily precipitation in QTP.

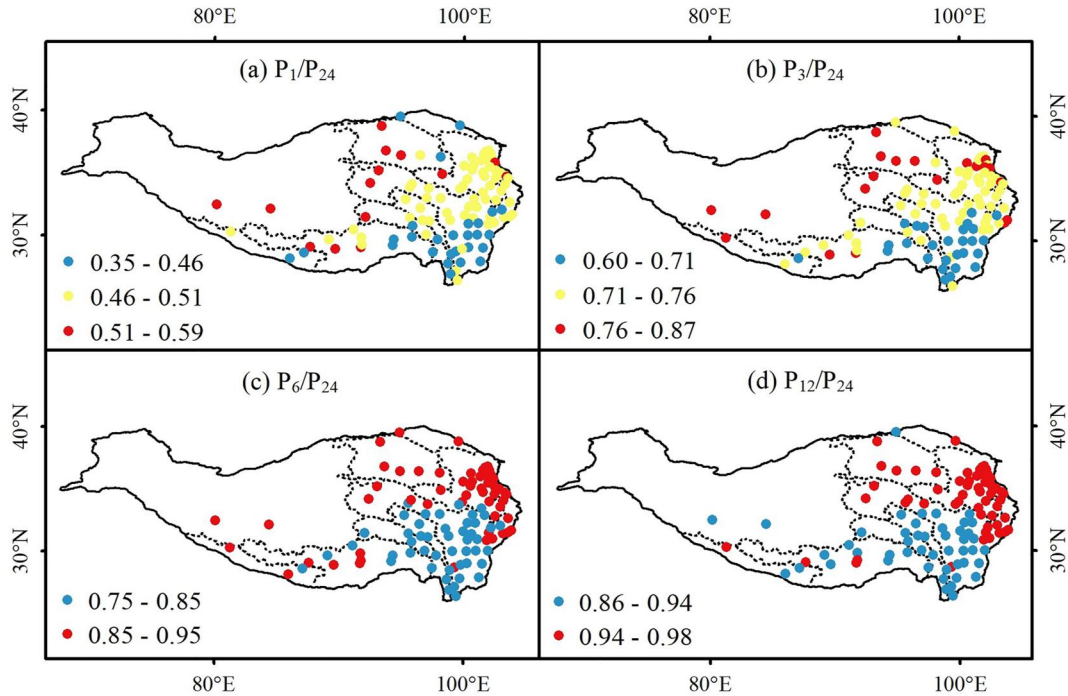

Figure 2. Ratios of sub-daily precipitation (t = 1, 3, 6, 12 hr) to daily precipitation at the 102 weather stations in the Qinghai-Tibet Plateau.

2.2.2. Geographical and Climatic Data

The digital elevation model (DEM) data, with a 1-km resolution, was used to extract the values of the geographical variables including longitude, latitude, and altitude in the study area. The DEM data were obtained from the United States Geological Survey. Besides, the daily data of precipitation, surface temperature and relative air humidity were obtained from CMA, and their mean values were calculated. The three geographical variables and three climatic variables are used to further investigate their influences on the temporal scaling characteristics of sub-daily precipitation in QTP.

3. Methodology

3.1. Fitting of $P_t/P_{24} \sim t$ Curve

In this study, the linear, logarithmic, exponential, power and inverse functions are used and compared, to choose the most suitable function for describing the temporal scaling characteristics of sub-daily precipitation in QTP. These functions are shown in Table 1.

Table 1
Types of Functions Used in This Study for Fitting the $P_t/P_{24} \sim$ t Curve[a]

Type	Equation
Linear function	$f(t) = a * t + b$
Logarithmic function	$f(t) = a * \ln t + b$
Exponential function	$f(t) = a * e^{bt}$
Power function	$f(t) = a * t^b$
Inverse function	$f(t) = a * (1/t) + b$

[a]Here $f(t)$ is P_t/P_{24}, and t is hour order, with t = 1, 2, …, 12 hr concerned in this study; a and b are two key parameters in the corresponding function.

The indicator of coefficient of determination (R^2) is used to evaluate the goodness of fit of these functions:

$$R^2 = 1 - \frac{\sum\left(f(t) - f(t)'\right)^2}{\sum\left(f(t) - \overline{f(t)}\right)^2}, \quad (1)$$

where $f(t)$ is the average $P_t/P_{24} \sim t$ curve calculated from the observed data at a station, and $\overline{f(t)}$ is the mean of $f(t)$; $f(t)'$ is the corresponding $P_t/P_{24} \sim t$ curve estimated by using certain fitting function in Table 1. R^2 with a value closer to 1 indicates a better fitting result by certain function. The p-value of 0.01 from the F-statistic test is further used to judge the statistical significance of the

fitting result at the 1% significance level. The fitting function with the largest R^2 value is chosen as the best one to describe the temporal scaling characteristics of sub-daily precipitation.

Further, $C_v{}^t$ and $C_s{}^t$ at each t-duration (with t = 1, 2, …, 12 hr) are calculated as:

$$C_v{}^t = \frac{\sigma}{f(t)} = \sqrt{\frac{\sum_{i=1}^{n}\left(f_i(t) - f(t)\right)^2}{f(t)^2(n-1)}} \tag{2}$$

$$C_s{}^t = \frac{\sum_{i=1}^{n}\left(f_i(t) - f(t)\right)^3}{f(t)^3(n-1)\,C_v{}^{t3}}, \tag{3}$$

where $f_i(t)$ is the $P_t/P_{24} \sim t$ curve calculated from the observed data in ith day, with the total number n of wet days, and $f(t)$ is its average $P_t/P_{24} \sim t$ curve. Note that these weather stations have different number n of wet days. Based on it, the $C_v{}^t \sim t$ curve and $C_s{}^t \sim t$ curve at each station can be obtained to quantify the uncertainty of the average $P_t/P_{24} \sim t$ curve.

3.2. Explanation of $P_t/P_{24} \sim t$ Curve

Considering the potential influences of geographical and climatic conditions on the temporal scaling characteristics of sub-daily precipitation, three geographical variables (longitude, latitude and altitude) and three climatic variables (average daily precipitation, surface temperature and relative air humidity) are used to explain the two key parameters in the chosen function. Here, the multiple linear regression (MLR) model, support vector regression (SVR) model and random forest regression (RFR) model, which are widely used for exploring the connections among variables (Chen et al., 2017; Ganguli & Reddy, 2014; Sachindra et al., 2018; Sajadi et al., 2022; Yan et al., 2022), are applied to establish the relationship between the six explanatory variables and the two key parameters. It is noted that the three models have been widely applied for spatial interpolation of monthly and daily precipitation in QTP (He et al., 2023; Jing et al., 2016; Z. Ma et al., 2018; Shi & Song, 2015), while their suitability and applicability for handling spatial interpolation of hourly precipitation in the region will be investigated in this study. The details of the three models can be found in Feidas et al. (2014) and Rodriguez-Galiano et al. (2015). Moreover, the normalized Taylor diagram, with three statistics of normalized standard deviation (σ_{nor}), normalized centered root-mean-square error ($CRMSE_{nor}$), and Pearson correlation coefficient (r), is used to evaluate the performance of the three models (Kärnä & Baptista, 2016; Taylor, 2001).

In the modeling practice, the data at 68 weather stations (accounting for 67% of the total stations) are used for calibration, and the data at the residual 34 weather stations (accounting for 33% of the total stations) are used for validation. Based on the results in the normalized Taylor diagram, the most suitable model can be chosen and established. After that, the data of the geographical and climatic variables in each 1-km grid are used as inputs to run the established model, to obtain the two key parameters in all grids. Based on it, the 1-km gridded parameters data set covering the whole QTP can be generated.

4. Results

4.1. Temporal Scaling Characteristics of Sub-Daily Precipitation

Taking a station (with the star symbol shown in Figure 1) as an example, its $P_t/P_{24} \sim t$ curves in all wet days are calculated and described by using boxplots, with its average $P_t/P_{24} \sim t$ curve shown as blue triangle points in Figure 3a. It presents that the average P_t/P_{24} value keeps increasing with t. Moreover, the dispersible characteristics (i.e., box length in Figure 3a and quantified by $C_v{}^t$ in Figure 3b) of P_t/P_{24} with smaller t-duration is bigger, implying larger uncertainty. Along with the increase of t-duration, the dispersible characteristics of P_t/P_{24} become smaller, corresponding to decreasing $C_v{}^t$ value, but P_t/P_{24} indicates more obviously negative skew characteristics (quantified by $C_s{}^t$ in Figure 3b).

The average $P_t/P_{24} \sim t$ curve is then fitted by using the five functions in Table 1. Figure 3a shows obvious difference in their fitting performances. The fitting curves by using the linear function and exponential function

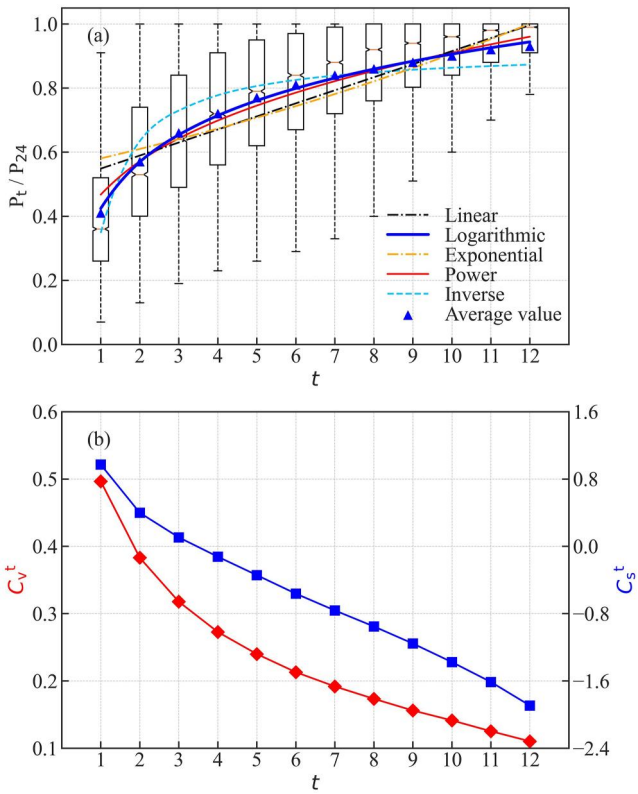

Figure 3. (a) Boxplots of the $P_t/P_{24} \sim t$ curves at an example station and their average value (blue triangle points), and the fitting curves of five functions; and (b) coefficient of variance (C_v^t), and coefficient of skewness (C_s^t) of the P_t/P_{24} (t = 1, 2, …, 12 hr) curves. The location of the example station (with the star symbol) is shown in Figure 1.

tend to be linear, which are far from the shape of the average $P_t/P_{24} \sim t$ curve. The other three functions provide similar fitting results, and the result from the logarithmic function is visually the best.

The performances of fitting results by using the five functions at the 102 stations are evaluated, and the results are shown in Figure 4a. It indicates that the logarithmic function provides the best fitting results, with an average R^2 value of 0.98 and a value range of 0.86–0.99. More importantly, the uncertainty of the fitting result by using the logarithmic function is the smallest (with the shortest box length in Figure 4a). Comparatively, the fitting results of the inverse function and power function have the average R^2 value of 0.95 and 0.94, respectively, which are a little worse than that of the logarithmic function. However, the linear function and exponential function performed badly and cannot catch the shape of the $P_t/P_{24} \sim t$ curve. It proves the nonlinear relationship between daily and sub-daily precipitation in QTP, just as shown in Figure 3a. While the approach of exploring the linear scaling relationship between daily and hourly precipitation, as done in previous studies (Cattoën et al., 2020; Lee & Jeong, 2014; Lee & Park, 2017), is not suitable for the study area.

Based on the above results, the logarithmic function of $P_t/P_{24} = a_{mean} \times \ln t + b_{mean}$ was chosen to describe the temporal scaling characteristics of sub-daily precipitation in QTP. In the function, parameter b_{mean} reflects the ratio between 1-hr and 24-hr precipitation, that is, $b_{mean} = P_1/P_{24}$; a larger a_{mean} means a higher P_t/P_{24} ratio at certain t-duration. Among the 102 stations, parameter a_{mean} has a value range of 0.13–0.21 and parameter b_{mean} has a value range of 0.37–0.66, as shown in Figures 5a and 5b, respectively. Regarding the spatial distribution, parameter a_{mean} (b_{mean}) has a southeast-northwest decreasing (increasing) pattern. The distributions of the two parameters reflect the spatial pattern of temporal scaling characteristics of sub-daily precipitation. Based on the physical meanings of the two parameters, it is known that heavy precipitation events with short durations usually occurred in western and central QTP with dry climatic conditions; comparatively, longer-duration precipitation more likely occurred in the southeastern QTP with wet climatic conditions. This spatial pattern is consistent with that in Figure 2, supporting the reliability of the chosen logarithmic function for describing the temporal scaling characteristics of sub-daily precipitation.

In addition, the $C_v^t \sim t$ curves and $C_s^t \sim t$ curves at the 102 weather stations are also fitted by using the five functions. Results show that the logarithmic function ($C_v^t = a_{cv} \times \ln t + b_{cv}$) gives the best fitting results for the $C_v^t \sim t$ curves (Figure 4b), and the linear function ($C_v^t = a_{cs} \times t + b_{cs}$) provides the best fitting results for the $C_s^t \sim t$ curves (Figure 4c). The parameters a_{cv}, b_{cv}, a_{cs}, and b_{cs} at the 102 stations have the value ranges of −0.15–−0.10, 0.33–0.53, −0.41–−0.14, −0.21–0.92, respectively, with their spatial distributions shown in Figures 5c–5f. Larger absolute values of parameters a_{cv} and b_{cv} in south part of QTP imply bigger dispersible characteristics of sub-daily precipitation compared to that in the northern part. Besides, larger absolute values of parameter a_{cs} (i.e., the slope of the chosen linear function) in the north reflect more obviously skew characteristics of sub-daily precipitation compared to that in the south part.

4.2. Relationship Between $P_t/P_{24} \sim t$ Curve and Explanatory Variables

The three geographical variables and three climatic variables are used to explain the key parameters of a_{mean} and b_{mean} in the chosen logarithmic function and their spatial distribution in Figure 5, to investigate the geographical and climatic influences on the spatial pattern of the temporal scaling characteristics of sub-daily precipitation in QTP. These explanatory variables can basically reflect both the moisture availability and redistribution of water vapor, which to a large extent affect precipitation amount and its spatial distribution (Jones et al., 2010; Lepore et al., 2016; Li et al., 2021). Figure S1 in Supporting Information S1 shows the different correlation degrees and directions of these explanatory variables with a_{mean} and b_{mean}, reflecting their different effects.

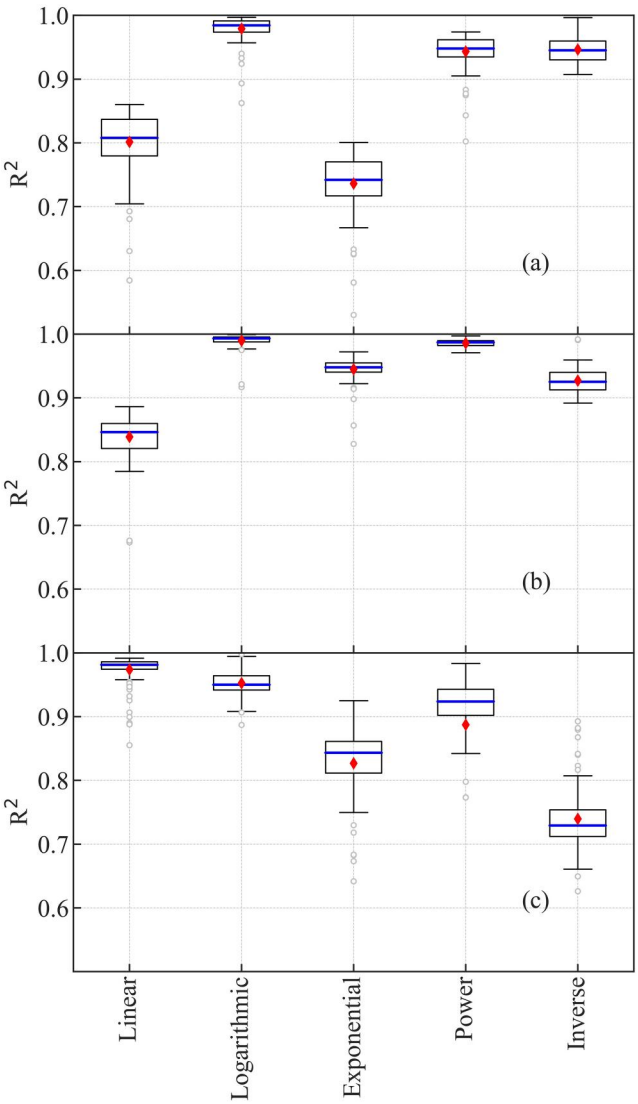

Figure 4. Boxplots of R^2 values of the five functions for fitting the $P_t/P_{24} \sim t$ curves (a), $C_v^t \sim t$ curves (b) and $C_s^t \sim t$ curves (c) at the 102 weather stations considered in this study. The blue solid line and red diamond point refer to the median and mean value of R^2, respectively.

Three models (MLR, SVR, RFR) are used to establish the relationship between the explanatory variables and each parameter. To explore the different effects of the geographical and climatic conditions on the two parameters, three types of explanatory variables are set here: three geographical variables (denoted as "geo"), three climatic variables (denoted as "clim"), and all the six variables (denoted as "geoclim"). The modeling results are evaluated by the normalized Taylor diagram and are shown in Figure 6. In this diagram, the evaluation metric of modeling result being closer to the original value (the black point) indicates a better accuracy (Xiao et al., 2020). It shows that no matter which types of explanatory variables are used as inputs, the MLR model performed badly, indicating that the relationships between the two parameters and the six explanatory variables cannot be explained well by the linear model, implying their nonlinear relationships. Comparatively, the SVR model performed well. For parameter a_{mean}, the value of r and $CRMSE_{nor}$ are 0.82 and 0.58 in the SVR(geo) model. For parameter b_{mean}, the value of r and $CRMSE_{nor}$ are 0.86 and 0.51 in the SVR(geoclim) model. Besides, the RFR model performed worse than SVR, which may be due to the small data sample used here, as the former usually requires a large sample size to ensure accurate modeling results (Millard & Richardson, 2015). Comparatively, the SVR model has a high generalization capability, particularly with respect to limited samples (Shao & Lunetta, 2012), which may explain its better performance compared to RFR.

Furthermore, Figure 6 indicates the different effects of the geographical and climatic variables. For parameter a_{mean}, the SVR(geo) model performs better than SVR(clim) model and SVR(geoclim) model (Figure 6a), suggesting that climate variables cannot give more useful explanations for the spatial pattern of a_{mean}. While for parameter b_{mean}, the SVR(geoclim) model outperforms SVR(geo) model and SVR(clim) model (Figure 6b). Although the three climatic variables contribute a part to the modeling results of parameter b_{mean}, it should be noticed that their values may have uncertainty due to the short data periods used in some weather stations. Although some other satellite- or reanalysis-based data sources can be used to estimate the values of the three climatic variables, their accuracy may have big biases in QTP (Hu et al., 2019; Simmons et al., 2010; J. Zhang et al., 2021), and their spatial resolution is much coarser than 1-km resolution concerned in this study. Comparatively, the values of the three geographical variables are stable, and the geographical conditions can to some extent reflect the climatic conditions. The SVR(geo) model with the three geographic variables gives the r of 0.77, σ_{nor} of 0.71 and $CRMSE_{nor}$ of 0.64 for parameter b_{mean}, whose accuracy is thought acceptable in such an ungauged region of QTP.

By considering the different performances of the three models, as well as the different influences of these explanatory variables and their data availability and reliability, the SVR model is selected, and the geographic variables of longitude, latitude, and altitude at 1-km grid are used as inputs, based on which the 1-km gridded parameter a_{mean} and b_{mean} in the whole QTP are generated, as shown in Figure 7. The value range of a_{mean} and b_{mean} in the whole QTP is from 0.14 to 0.20 and from 0.38 to 0.65, respectively. For parameter a_{mean}, its high values appear in the eastern, southeastern, and southern boundary areas of QTP (Figure 7a), where natural disasters induced by rainstorm events occurred frequently (Wang et al., 2021). While the value of a_{mean} is relatively low in central QTP, including the Qaidam Basin, the headwater regions of the Jinsha River Basin, the Mekong River Basin, the Salween River Basin, and the mid part of the Yarlung Tsangpo River Basin. For parameter b_{mean}, its high values appear in the Qaidam Basin and small values in southeast QTP (Figure 7b). As a result, the two parameters of a_{mean} and b_{mean} can determine the logarithmic $P_t/P_{24} \sim t$ curve at each grid, as a basis of estimating t-hour precipitation, especially for those areas with only daily precipitation records available.

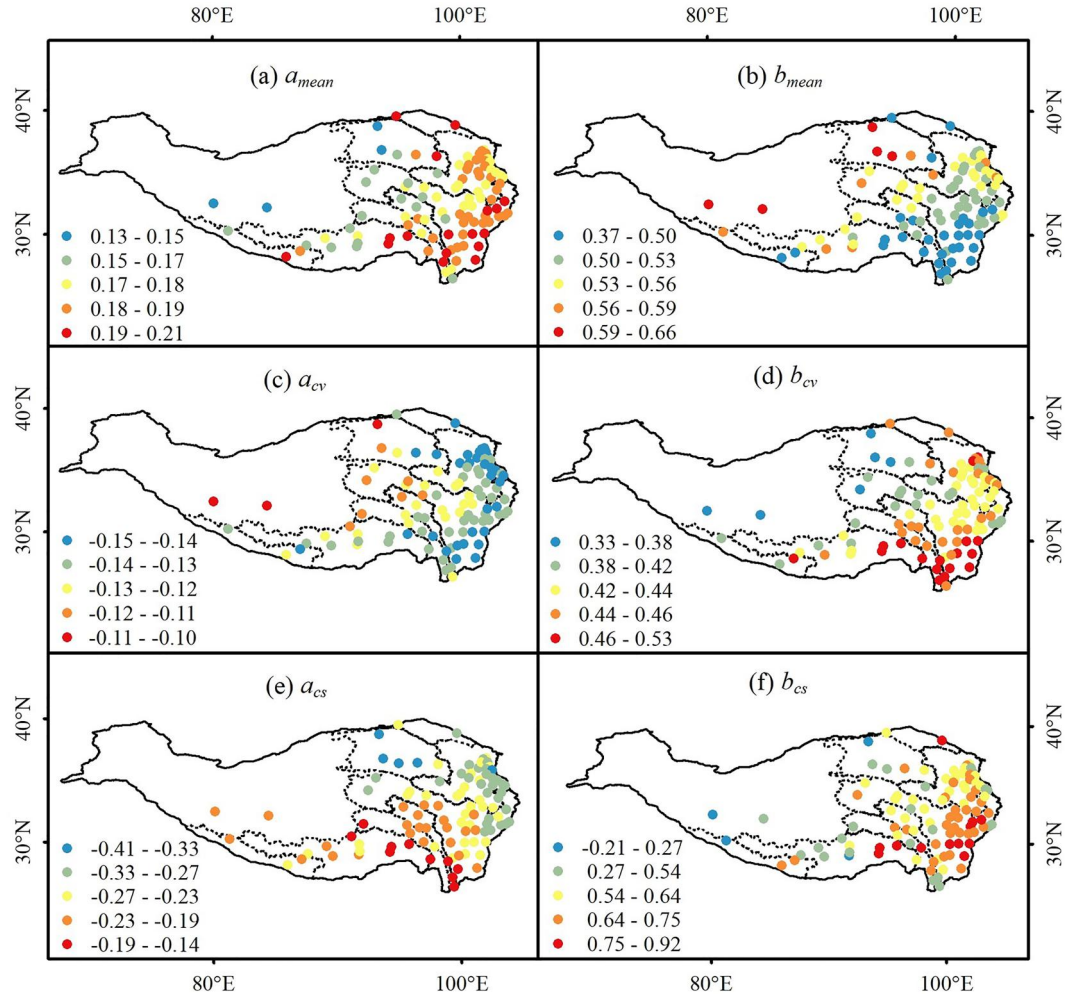

Figure 5. Spatial distributions of parameters a_{mean} (a), b_{mean} (b), a_{cv} (c) b_{cv} (d), a_{cs} (e), and b_{cs} (f) at the 102 weather stations in Qinghai-Tibet Plateau (QTP). Regarding these parameters, the logarithmic function of $\mathbf{P_t/P_{24} = a_{mean} \times \ln t + b_{mean}}$ is chosen to describe the $P_t/P_{24} \sim t$ curve; the logarithmic function of $\boldsymbol{C_v^t = a_{cv} \times \ln t + b_{cv}}$ is chosen to describe the $C_v^{\,t} \sim t$ curve; and the linear function of $\boldsymbol{C_s^t = a_{cs} \times t + b_{cs}}$ is chosen to describe the $C_s^{\,t} \sim t$ curve.

4.3. Uncertainty Quantification

For quantifying the uncertainty of the logarithmic $P_t/P_{24} \sim t$ curve, we estimated the logarithmic $C_v^{\,t} \sim t$ curve and the linear $C_s^{\,t} \sim t$ curve (as shown in Figure 3b) by using the MLR, SVR, and RFR model, and the three types of explanatory variables are used again. The correlations of these parameters with the six explanatory variables can be found in Figures S2 and S3 of the Supporting Information S1. The results in Figure S4 of the Supporting Information S1 present the different performances of the three models, among which the SVR model performed well, being similar to those in Figure 6.

By considering the data availability and reliability again, the SVR model with the inputs of three geographical variables is chosen to estimate the key parameters in the logarithmic $C_v^{\,t} \sim t$ curve and in the linear $C_s^{\,t} \sim t$ curve, based on which the 1-km gridded parameters of a_{cv}, b_{cv}, a_{cs}, and b_{cs} in the whole QTP are generated, and their spatial distribution are shown in Figure 8. For the two parameters a_{cv} and b_{cv} in logarithmic $C_v^{\,t} \sim t$ curve, their value ranges are −0.16–−0.10 and 0.35–0.52, respectively. High values of a_{cv} appear in the Qaidam Basin and central QTP, and low values of a_{cv} appear in the eastern, southeastern, and southern boundary areas of QTP (Figure 8a). The spatial distribution of b_{cv} is opposite of that of a_{cv}, with low values appearing in the Qaidam Basin and high values in eastern, southeastern and southern QTP. For the two parameters a_{cs} and b_{cs} in linear $C_s^{\,t} \sim t$ curve, their values ranges are −0.41–−0.13 and −0.17–0.88, respectively. High values of a_{cs} appear in the

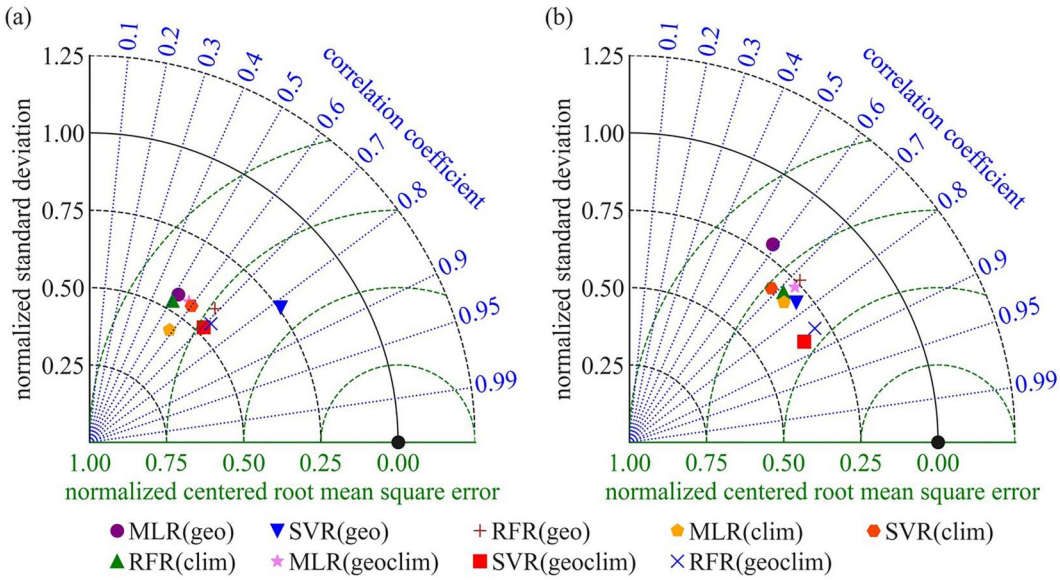

Figure 6. Normalized Taylor diagrams showing the performances of the MLR, SVR, RFR models for estimating parameter a_{mean} (a) and b_{mean} (b). Normalized standard deviation (σ_{nor}) is on the radial axis; correlation coefficient (r) is on the angular axis; and green dashed lines indicate $CRMSE_{nor}$. The suffixes "geo," "clim", and "geoclim" refer to models with the input data of geographical variables, climatic variables, and a combination of them, respectively. The parameters a_{mean} and b_{mean} are the same as that in Figure 5.

northeast corner and the south part of the study area, while low values of a_{cs} appear in the north part. Regarding parameter b_{cs}, its high values appear in the southeastern QTP, including the Yellow River Basin, the origins of the Jinsha River Basin, the Mekong River Basin, the Salween River Basin, and the outlet of the Yarlung Tsangpo River Basin; and low values of b_{cs} appear in central QTP and the Qaidam Basin.

By using the estimated four parameters in Figure 8, the probabilistic distribution of the P_t/P_{24} value at any t-durations in each grid can be mathematically described, and it can be a reliable basis of quantitative uncertainty evaluation of the sub-daily precipitation (i.e., P_t) estimated from Figure 7, especially for short time-duration with high uncertainty (as the example shown in Figure 3). Therefore, the 1-km gridded data set of the six key parameters can catch the essential information of the complex extreme precipitation conditions over QTP.

5. Summary and Discussion

5.1. Main Contributions of This Study

QTP is a typical ungauged region, and the precipitation data at sub-daily scales for the studies of natural disasters are limited. For filling the data gap of sub-daily precipitation in QTP, this study developed a new strategy to explore the temporal scaling characteristics of sub-daily precipitation in the region, by considering the influences of geographical and climatic conditions. Based on it, a 1-km gridded parameters data set for estimating sub-daily precipitation in the whole QTP was generated. The results indicated that a logarithmic function with two parameters of a_{mean} and b_{mean} can well describe the temporal scaling characteristics of sub-daily precipitation (denoted as $P_t/P_{24} \sim t$ curve) at the 102 weather stations in QTP. Besides, the logarithmic $C_v^t \sim t$ curve with the parameters of a_{cv} and b_{cv}, and linear $C_s^t \sim t$ curve with the parameters of a_{cs} and b_{cs}, can quantitatively evaluate the uncertainty of the $P_t/P_{24} \sim t$ curve. Moreover, each of the above six parameters is closely related to the geographical conditions, and their relationships can be well established by the SVR model. By using the model, three geographic variables of longitude, latitude, and altitude at 1-km grid are used as inputs, based on which the 1-km gridded data set for the six parameters in the whole QTP is generated. The parameters data set reflects the essential information of the complex extreme precipitation conditions over QTP, and thus it can provide a feasible approach for estimating sub-daily precipitation, supporting the studies of rainstorm-related natural disasters in the region.

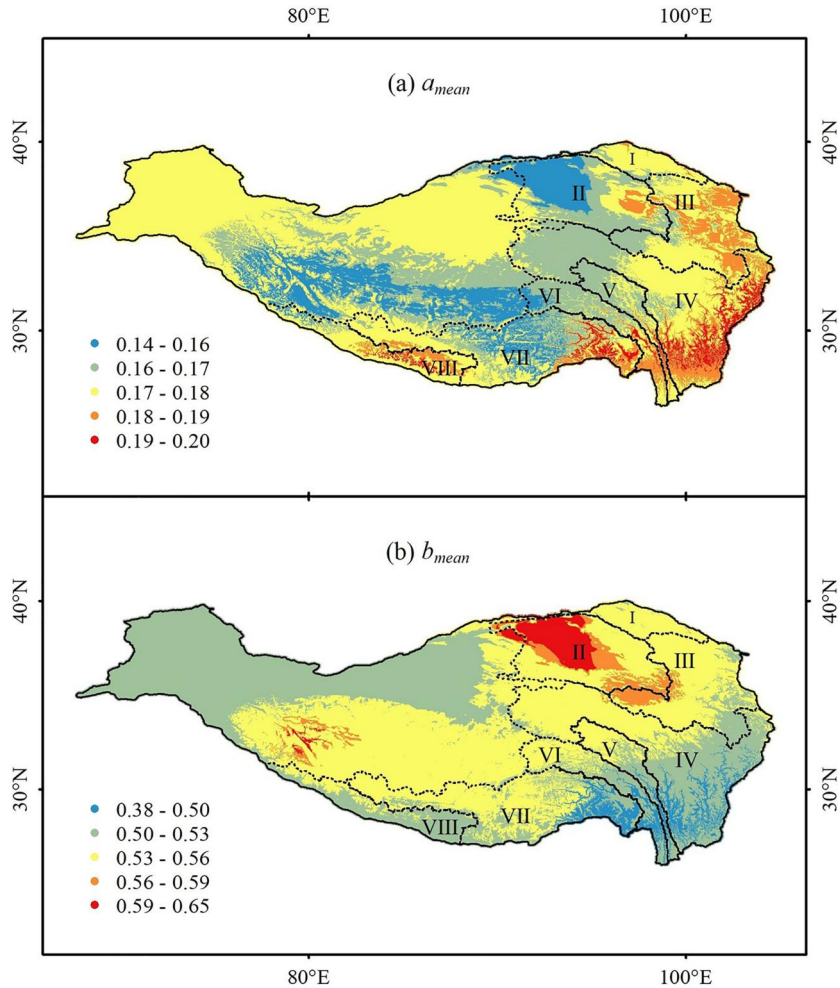

Figure 7. 1-km gridded data set of the parameters of a_{mean} (a) and b_{mean} (b) over QTP. The two parameters determine the logarithmic $P_t/P_{24} \sim t$ curve at each grid, and they were generated by using the SVR model, with the geographic variables of longitude, latitude, and altitude as inputs. The eight basins are the same as those in Figure 1.

Some satellite data and reanalysis data of daily precipitation covering the whole QTP are available. Based on them, the maximum t-duration precipitation (t = 1, 2, …, 12 hr) in local areas can be estimated by using the logarithmic $P_t/P_{24} \sim t$ curve established. Particularly, the probabilistic distribution of the maximum t-duration precipitation can also be quantitatively described by using the logarithmic $C_v^t \sim t$ curve and linear $C_s^t \sim t$ curve established. Following the probabilistic distribution, it would be favorable and flexible for extracting the key t-duration precipitation information at any concerned significance levels (or confidence intervals) by researchers and engineers, as a necessary basis for risk analysis and also for forecasting and warning of rainstorm-related natural disasters in QTP.

It should be emphasized that this study revealed the simple and smooth temporal scaling characteristics of sub-daily precipitation (see Figure 3a), as well as their regional similarities (see Figures 2, 5, 7, and 8) due to the effects of geographical and climatic conditions in QTP. Thus, the strategy developed in this study to explore the describable temporal scaling characteristics of sub-daily precipitation is thought as an innovative approach for temporal downscaling, as it is simpler, easier to understand, and more effective compared to those conventional methods which tried to establish the complex and indescribable relationship between daily and sub-daily precipitation (Müller & Haberlandt, 2018; Socolofsky et al., 2001). This innovative strategy is recommended for estimating sub-daily precipitation and evaluating its uncertainty in those regions with the lack of short-duration precipitation data.

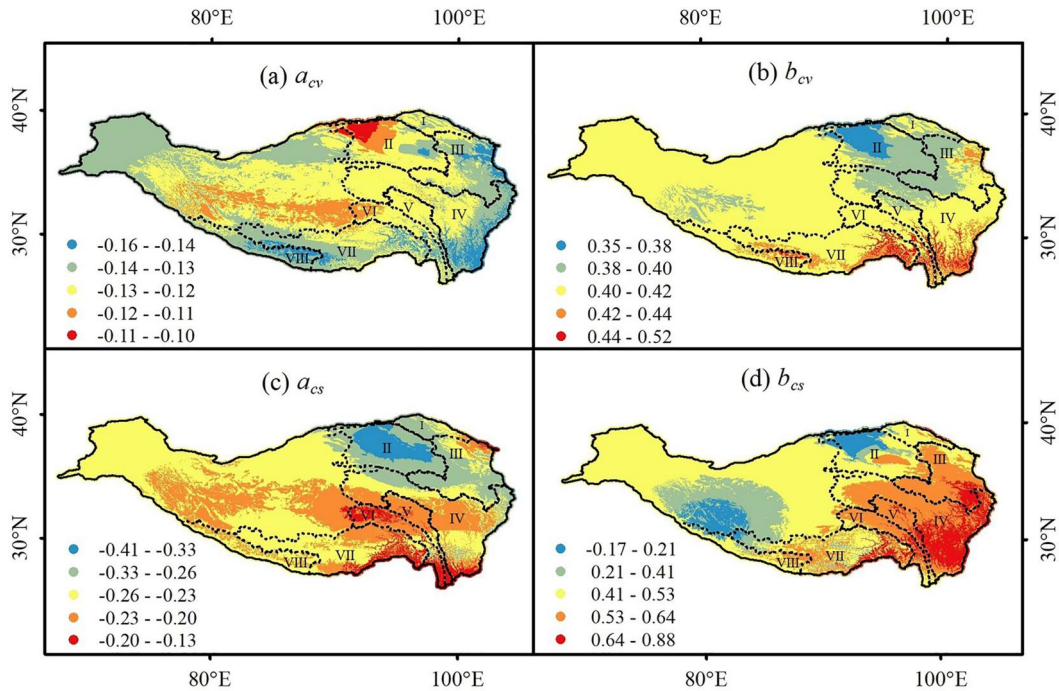

Figure 8. Spatial distribution of the parameters a_{cv} (a), b_{cv} (b), a_{cs} (c), and b_{cs} (d) in 1-km grids in QTP. The four parameters are the same as those in Figure 5. They were generated by using the SVR model, with the geographic variables of longitude, latitude, and altitude as inputs. The eight basins are the same as those in Figure 1.

The potential advantages of the strategy developed in this study are explained as follows. It could be taken as a new criterion for evaluating the quality of alternative sources (satellite-based or reanalysis-based) of hourly precipitation data. The strategy should be more effective and reliable compared to those conventional indicators such as R^2, σ_{nor}, $CRMSE_{nor}$, and r, as they can only describe the simple statistical relationship between the alternative data sources and field observed hourly data, but cannot quantify the temporal scaling characteristics of sub-daily precipitation. Besides, the temporal scaling characteristics of sub-daily precipitation can be further employed as useful guidance for bias correction of these alternative hourly precipitation data sources. Finally, if the temporal scaling characteristics of sub-daily precipitation is supposed to keep stable and does not change under climate change, the strategy can also be used to handle the climatic model data under different climate scenarios, as the extreme precipitation at sub-daily scales (but not only at daily scales) must be considered for the projection of future risks of rainstorm-related natural disasters and the climate change adaption in high mountain regions.

5.2. The Uncertainty Associated With the Parameters Data Set

The reliability of the parameters data set generated in this study could be influenced by two sources of uncertainty: the quantity of field hourly precipitation observations and the selection of explanatory variables. In our study, we incorporated all the available hourly precipitation data from the 102 weather stations, ensuring that there was no overlap in their data periods. This approach was implemented to capture the maximum amount of relevant statistical characteristics of hourly precipitation and to utilize the available information to the fullest extent. The hourly precipitation data records are longer than 10 years at 98 stations, and longer than 30 years at 38 stations. The different data lengths may influence the stability of the $P_t/P_{24} \sim t$ curves. To clarify it, we used the same example station as that in Figure 3, and calculated the $P_t/P_{24} \sim t$ curves corresponding to six data periods. Unexpectedly, Figure 9 shows that the $P_t/P_{24} \sim t$ curves are consistent despite different data records being used. It may be due to the intrinsic stability and consistency of the temporal scaling characteristics of sub-daily precipitation, which is controlled by its specific geographic and climatic conditions. Thus, it is thought that the data lengths have a weak influence on the stability of the $P_t/P_{24} \sim t$ curves. This stability on the one hand proves the

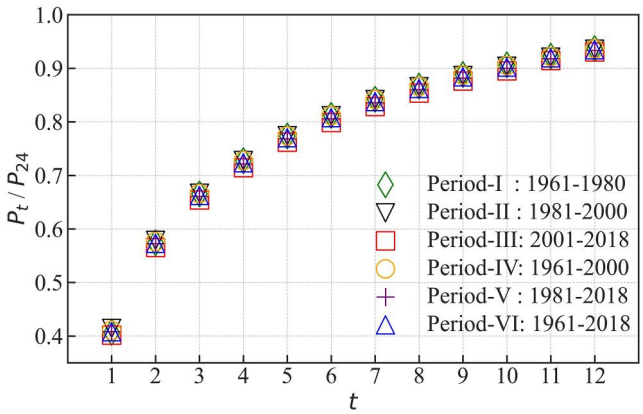

Figure 9. The average $P_t/P_{24} \sim t$ curves calculated by using six data periods at the example station, with its location (star symbol) shown in Figure 1.

advantage of the innovative strategy developed in this study, on the other hand, it reflects the reliability of the parameters data set generated in Figures 7 and 8.

The number of weather stations may also impact the reliability of the parameters data set generated. In this study, 100 stations are located in the river valley region of the eight basins, based on which the parameters data set generated in the region could be reliable. However, the central and western parts of QTP only include two weather stations (see Figure 1), where rainstorm-related natural disasters historically occurred much less frequently than in the river valley regions (Wang et al., 2021). To present the big picture, we used the SVR model calibrated in the river valley region to estimate the six parameters at each 1-km grid in the whole QTP. In the future, more field precipitation observations with long periods are still required to further verify the reliability of the parameters data generated, especially for the central and western parts of QTP.

Another issue is about the types of precipitation. Considering that rainstorm-related natural disasters in QTP mainly occur in summer, in this study we mainly analyzed the hourly rainfall data from May to September, without considering sleet and snow phenomena in those high areas above the snow line (You et al., 2020; Yuan et al., 2022). However, it should be noticed that the population and economy are mainly distributed below and far from the snow line. Thus, the issue would not influence the practical application of the parameters data set generated for the studies of rainstorm-related natural disasters in QTP.

Moreover, in this study, we mainly used the geographical variables of longitude, latitude and altitude to explain the six key parameters. As explained in Section 4.2, although the three climatic variables contribute a part to the modeling results of some parameters, their effects are much weaker than those of the three geographical variables. More importantly, high spatial resolution (1-km) data of the three climatic variables are unavailable to support the generation of 1-km gridded parameters data set. Alternatively, the modeling accuracy (see Figure 6) based on the three geographical variables is thought to be acceptable, based on which the parameters data set generated in Figures 7 and 8 are thought to be reliable. In the future, more explanatory variables can be used to improve the quantification results of temporal scaling characteristics of sub-daily precipitation in the study area.

Besides, due to the limitation of available weather stations, some related issues such as the effects of altitudinal precipitation gradient (Jiang et al., 2022) and diurnal characteristics of sub-daily precipitation (Li, 2018; Zipser & Xu, 2011), cannot be fully taken into account in this study. Future studies considering these could further improve the quality of the parameters data set generated. Finally, it is pointed out that the durations of historical extreme rainfall and flooding disasters in QTP are from minutes to hours, and thus the precipitation information at sub-hourly scales is another open issue that should be explored in the future.

6. Conclusions

The Qinghai-Tibet Plateau (QTP) is known for its high vulnerability to destructive rainstorm hazards and related natural disasters. However, the study of these natural disasters in the region has been hindered by limited observed short-duration precipitation data. To address this data gap, we conducted an investigation into the temporal scaling characteristics of sub-daily precipitation in the QTP, and we further examined the influence of geographical and climatic conditions on this temporal scaling characteristics of sub-daily precipitation. As a result of our research, we have generated a comprehensive 1-km gridded parameter data set that can be used to estimate sub-daily precipitation throughout the entire QTP. The data set includes six key parameters, among which the first two parameters of a_{mean} and b_{mean} determine the temporal scaling characteristics of sub-daily precipitation described by a logarithmic $P_t/P_{24} \sim t$ curve, and the last four parameters (a_{cv}, b_{cv}, a_{cs}, b_{cs}) quantify the uncertainty of the temporal scaling characteristics of sub-daily precipitation. Overall, the parameters data set can quantitatively describe the probabilistic distribution of the maximum t-duration precipitation. Based on the probabilistic distribution, it can easily and flexibly extract the key t-duration precipitation information at any desired significance level, providing a fundamental basis for the risk analysis and forecasting/early warning of rainstorm-related natural disasters in QTP.

Finally, we would like to emphasize the significance of developing a quantifiable equation to describe the temporal scaling characteristics of sub-daily precipitation as an innovative approach to temporal downscaling. This methodology not only serves as a new and effective criterion for assessing the quality of various data sources but also enables the evaluation of bias correlation. To further validate the application and effectiveness of the innovative strategy devised in this study, it is recommended to conduct additional case studies in different regions with diverse geographical and climatic conditions.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The boundary data for QTP are available in Y. Zhang (2019), and the boundary data of eight basins are available in G. Zhang (2019). The DEM data are available through the website https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-30-arc-second-elevation-gtopo30. The generated parameters data set is publicly available in Sang (2024) and http://rcg.gvc.gu.se/. Python code used for producing this parameters data set is publicly available in Ren (2024).

Acknowledgments
This work is financially supported by the National Key Research and Development Program (Grant 2019YFA0606903), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (Grant 2019QZKK0906), and the National Natural Science Foundation of China (Grants 41971040, 42311530063). DC is supported by the Swedish Research Council (2021-02163).

References

Benestad, R. E., Chen, D., & Hanssen-Bauer, I. (2008). *Empirical-statistical downscaling*. World Scientific Publishing Company. https://doi.org/10.1142/6908

Benestad, R. E., Lutz, J., Dyrrdal, A. V., Haugen, J. E., Parding, K. M., & Dobler, A. (2021). Testing a simple formula for calculating approximate intensity-duration-frequency curves. *Environmental Research Letters*, *16*(4), 044009. https://doi.org/10.1088/1748-9326/abd4ab

Benestad, R. E., Nychka, D., & Mearns, L. O. (2012). Spatially and temporally consistent prediction of heavy precipitation from mean values. *Nature Climate Change*, *2*(7), 544–547. https://doi.org/10.1038/nclimate1497

Casanueva, A., Herrera, S., Iturbide, M., Lange, S., Jury, M., Dosio, A., et al. (2020). Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmospheric Science Letters*, *21*(7). https://doi.org/10.1002/asl.978

Cattoën, C., Robertson, D. E., Bennett, J. C., Wang, Q. J., & Carey-Smith, T. K. (2020). Calibrating hourly precipitation forecasts with daily observations. *Journal of Hydrometeorology*, *21*(7), 1655–1673. https://doi.org/10.1175/JHM-D-19-0246.s1

Chen, D., Fan, L., & Fu, C. (2005). Review on creating future climate change scenarios by statistical downscaling techniques. *Advances in Earth Science*, *20*, 320–329. https://doi.org/10.11867/J.ISSN.1001-8166.2005.03.0320

Chen, H., Hou, Y.-K., Xu, C.-Y., Chen, J., & Guo, S.-L. (2017). Coupling a Markov chain and support vector machine for at-site downscaling of daily precipitation. *Journal of Hydrometeorology*, *18*(9), 2385–2406. https://doi.org/10.1175/jhm-d-16-0130.1

Cui, P., & Jia, Y. (2015). Mountain hazards in the Tibetan Plateau: Research status and prospects. *National Science Review*, *2*(4), 397–399. https://doi.org/10.1093/nsr/nwv061

Feidas, H., Karagiannidis, A., Keppas, S., Vaitis, M., Kontos, T., Zanis, P., et al. (2014). Modeling and mapping temperature and precipitation climate data in Greece using topographical and geographical parameters. *Theoretical and Applied Climatology*, *118*(1), 133–146. https://doi.org/10.1007/s00704-013-1052-4

Ganguli, P., & Reddy, M. J. (2014). Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach. *Hydrological Processes*, *28*(19), 4989–5009. https://doi.org/10.1002/hyp.9966

Hanssen-Bauer, I., Achberger, C., Benestad, R. E., Chen, D., & Førland, E. J. (2005). Statistical downscaling of climate scenarios over Scandinavia. *Climate Research*, *29*, 255–268. https://doi.org/10.3354/cr029255

He, K., Chen, X., Liu, J., & Zhao, D. (2023). A multiple-step scheme for the improvement of satellite precipitation products over the Tibetan Plateau from multisource information. *Science of the Total Environment*, *873*, 162378. https://doi.org/10.1016/j.scitotenv.2023.162378

Hock, R. G., Rasul, C., Adler, B., Cáceres, S., Gruber, Y., Hirabayashi, M., et al. (2022). High mountain areas. In *The ocean and cryosphere in a changing climate* (pp. 131–202). https://doi.org/10.1017/9781009157964.004

Hu, G., Zhao, L., Wu, X., Li, R., Wu, T., Su, Y., & Hao, J. (2019). Evaluation of reanalysis air temperature products in permafrost regions on the Qinghai-Tibetan Plateau. *Theoretical and Applied Climatology*, *138*(3), 1457–1470. https://doi.org/10.1007/s00704-019-02888-8

Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., et al. (2019). Importance and vulnerability of the world's water towers. *Nature*, *577*(7790), 364–369. https://doi.org/10.1038/s41586-019-1822-y

Jiang, Y., Yang, K., Qi, Y., Zhou, X., He, J., Lu, H., et al. (2023). 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 System Science Data*, *15*(2), 621–638. https://doi.org/10.5194/essd-15-621-2023

Jiang, Y., Yang, K., Yang, H., Lu, H., Chen, Y., Zhou, X., et al. (2022). Characterizing basin-scale precipitation gradients in the Third Pole region using a high-resolution atmospheric simulation-based dataset. *Hydrology and Earth System Sciences*, *26*(17), 4587–4601. https://doi.org/10.5194/hess-26-4587-2022

Jing, W., Yang, Y., Yue, X., & Zhao, X. (2016). A spatial downscaling algorithm for satellite-based precipitation over the Tibetan plateau based on NDVI, DEM, and land surface temperature. *Remote Sensing*, *8*(8), 655. https://doi.org/10.3390/rs8080655

Jones, R. H., Westra, S., & Sharma, A. (2010). Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geographical Research Letters*, *37*(22), L22805. https://doi.org/10.1029/2010GL045081

Kärnä, T., & Baptista, A. M. (2016). Evaluation of a long-term hindcast simulation for the Columbia River estuary. *Ocean Modelling*, *99*, 1–14. https://doi.org/10.1016/j.ocemod.2015.12.007

Kukulies, J., Chen, D., & Wang, M. (2020). 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. *International Journal of Climatology*, *40*(11), 4858–4875. https://doi.org/10.1002/joc.6493

Lee, T., & Jeong, C. (2014). Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios. *Journal of Hydrology*, *510*, 182–196. https://doi.org/10.1016/j.jhydrol.2013.12.027

Lee, T., & Park, T. (2017). Nonparametric temporal downscaling with event-based population generating algorithm for RCM daily precipitation to hourly: Model development and performance evaluation. *Journal of Hydrology*, *547*, 498–516. https://doi.org/10.1016/j.jhydrol.2017.01.049

Lepore, C., Allen, J. T., & Tippett, M. K. (2016). Relationships between hourly rainfall intensity and atmospheric variables over the contiguous United States. *Journal of Climate*, *29*(9), 3181–3197. https://doi.org/10.1175/jcli-d-15-0331.1

Li, G., Yu, Z., Wang, W., Ju, Q., & Chen, X. (2021). Analysis of the spatial distribution of precipitation and topography with GPM data in the Tibetan Plateau. *Atmospheric Research*, *247*, 105259. https://doi.org/10.1016/j.atmosres.2020.105259

Li, J. (2018). Hourly station-based precipitation characteristics over the Tibetan Plateau. *International Journal of Climatology*, *38*(3), 1560–1570. https://doi.org/10.1002/joc.5281

Lima, C. H. R., Kwon, H.-H., & Kim, Y.-T. (2018). A local-regional scaling-invariant Bayesian GEV model for estimating rainfall IDF curves in a future climate. *Journal of Hydrology*, *566*, 73–88. https://doi.org/10.1016/j.jhydrol.2018.08.075

Liu, Z., Yao, Z., Wang, R., & Yu, G. (2020). Estimation of the Qinghai-Tibetan Plateau runoff and its contribution to large Asian rivers. *Science of the Total Environment*, *749*, 141570. https://doi.org/10.1016/j.scitotenv.2020.141570

Lu, D., & Yong, B. (2018). Evaluation and hydrological utility of the latest GPM IMERG V5 and GSMaP V7 precipitation products over the Tibetan Plateau. *Remote Sensing*, *10*(12), 2022. https://doi.org/10.3390/rs10122022

Ma, W., Liu, F., Zhou, Q., Chen, Q., Zhang, C., Liu, F., et al. (2021). Estimation of critical rainfall for flood disasters in the Qinghai-Tibet Plateau. *Journal of Resources and Ecology*, *12*(5), 600–608. https://doi.org/10.5814/j.issn.1674-764x.2021.05.003

Ma, Y., Yang, Y., Han, Z., Tang, G., Maguire, L., Chu, Z., & Hong, Y. (2018). Comprehensive evaluation of ensemble multi-satellite precipitation dataset using the dynamic Bayesian model averaging scheme over the Tibetan plateau. *Journal of Hydrology*, *556*, 634–644. https://doi.org/10.1016/j.jhydrol.2017.11.050

Ma, Z., He, K., Tan, X., Xu, J., Fang, W., He, Y., & Hong, Y. (2018). Comparisons of spatially downscaling TMPA and IMERG over the Tibetan Plateau. *Remote Sensing*, *10*(12), 1883. https://doi.org/10.3390/rs10121883

Ma, Z., Xu, J., Ma, Y., Zhu, S., He, K., Zhang, S., et al. (2022). AERA5-Asia: A long-term Asian precipitation dataset (0.1°, 1-hourly, 1951–2015, Asia) anchoring the ERA5-Land under the total volume control by APHRODITE. *Bulletin of the American Meteorological Society*, *103*(4), E1146–E1171. https://doi.org/10.1175/bams-d-20-0328.1

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., et al. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, *48*(3), RG3003. https://doi.org/10.1029/2009rg000314

Maskey, M. L., Puente, C. E., & Sivakumar, B. (2019). Temporal downscaling rainfall and streamflow records through a deterministic fractal geometric approach. *Journal of Hydrology*, *568*, 447–461. https://doi.org/10.1016/j.jhydrol.2018.09.014

Millard, K., & Richardson, M. (2015). On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sensing*, *7*(7), 8489–8515. https://doi.org/10.3390/rs70708489

MoradiKhaneghahi, M., Lee, T., & Singh, V. P. (2019). Stepwise extreme learning machine for statistical downscaling of daily maximum and minimum temperature. *Stochastic Environmental Research and Risk Assessment*, *33*(4–6), 1035–1056. https://doi.org/10.1007/s00477-019-01680-4

Müller, H., & Haberlandt, U. (2018). Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. *Journal of Hydrology*, *556*, 847–864. https://doi.org/10.1016/j.jhydrol.2016.01.031

Ou, T., Chen, D., Chen, X., Lin, C., Yang, K., Lai, H.-W., & Zhang, F. (2020). Simulation of summer precipitation diurnal cycles over the Tibetan Plateau at the gray-zone grid spacing for cumulus parameterization. *Climate Dynamics*, *54*(7–8), 3525–3539. https://doi.org/10.1007/s00382-020-05181-x

Parding, K. M., Benestad, R. E., Dyrrdal, A. V., & Lutz, J. (2023). A principal-component-based strategy for regionalisation of precipitation intensity–duration–frequency (IDF) statistics. *Hydrology and Earth System Sciences*, *27*(20), 3719–3732. https://doi.org/10.5194/hess-27-3719-2023

Pepin, N., Bradley, R. S., Diaz, H. F., Baraer, M., Caceres, E. B., Forsythe, N., & Mountain Research Initiative, E. D. W. W. G. (2015). Elevation-dependent warming in mountain regions of the world. *Nature Climate Change*, *5*(5), 424–430. https://doi.org/10.1038/nclimate2563

Pepin, N. C., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., et al. (2022). Climate changes and their elevational patterns in the mountains of the world. *Reviews of Geophysics*, *60*(1), e2020RG000730. https://doi.org/10.1029/2020rg000730

Ren, Z. (2024). 68kellyren/datacode (version 1.1) [Software]. Zenodo. https://doi.org/10.5281/zenodo.10477511

Rodriguez-Galiano, V. F., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, *71*, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Sachindra, D. A., Ahmed, K., Rashid, M. M., Shahid, S., & Perera, B. J. C. (2018). Statistical downscaling of precipitation using machine learning techniques. *Atmospheric Research*, *212*, 240–258. https://doi.org/10.1016/j.atmosres.2018.05.022

Sajadi, P., Sang, Y.-F., Gholamnia, M., Bonafoni, S., & Mukherjee, S. (2022). Evaluation of the landslide susceptibility and its spatial difference in the whole Qinghai-Tibetan Plateau region by five learning algorithms. *Geoscience Letters*, *9*(1), 9. https://doi.org/10.1186/s40562-022-00218-x

Sang, Y. (2024). A high-resolution gridded (1km×1km) parameters data set for estimating sub-daily precipitation across the Qinghai-Tibet Plateau [Dataset]. National Tibetan Plateau Data Center. https://doi.org/10.11888/Atmos.tpdc.300983

Shao, Y., & Lunetta, R. S. (2012). Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing*, *70*, 78–87. https://doi.org/10.1016/j.isprsjprs.2012.04.001

Shi, Y., & Song, L. (2015). Spatial downscaling of monthly TRMM precipitation based on EVI and other geospatial variables over the Tibetan Plateau from 2001 to 2012. *Mountain Research and Development*, *35*(2), 180–194. https://doi.org/10.1659/MRD-JOURNAL-D-14-00119.1

Simmons, A. J., Willett, K. M., Jones, P. D., Thorne, P. W., & Dee, D. P. (2010). Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets. *Journal of Geophysical Research*, *115*(D1), D01110. https://doi.org/10.1029/2009JD012442

Socolofsky, S., Adams, E. E., & Entekhabi, D. (2001). Disaggregation of daily rainfall for continuous watershed modeling. *Journal of Hydrologic Engineering*, *6*(4), 300–309. https://doi.org/10.1061/(asce)1084-0699(20016:4(300))

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017rg000574

Sun, S., Zhang, Q., Xu, Y., & Yuan, R. (2021). Integrated assessments of meteorological hazards across the Qinghai-Tibet Plateau of China. *Sustainability*, *13*(18), 10402. https://doi.org/10.3390/su131810402

Takhellambam, B. S., Srivastava, P., Lamba, J., McGehee, R. P., Kumar, H., & Tian, D. (2022). Temporal disaggregation of hourly precipitation under changing climate over the Southeast United States. *Scientific Data*, *9*(1), 211. https://doi.org/10.1038/s41597-022-01304-7

Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA version-7 legacy products over Mainland China at multiple spatiotemporal scales. *Journal of Hydrology*, *533*, 152–167. https://doi.org/10.1016/j.jhydrol.2015.12.008

Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., & Wu, J. (2016). Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *Journal of Geophysical Research: Atmospheres*, *121*(5), 2110–2129. https://doi.org/10.1002/2015jd023977

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000jd900719

Terzi, S., Torresan, S., Schneiderbauer, S., Critto, A., Zebisch, M., & Marcomini, A. (2019). Multi-risk assessment in mountain regions: A review of modelling approaches for climate change adaptation. *Journal of Environmental Management*, *232*, 759–771. https://doi.org/10.1016/j.jenvman.2018.11.100

Tong, K., Su, F., Yang, D., & Hao, Z. (2014). Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *Journal of Hydrology*, *519*, 423–437. https://doi.org/10.1016/j.jhydrol.2014.07.044

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*(7), W07447. https://doi.org/10.1029/2006wr005653

Vu, M. T., Aribarg, T., Supratid, S., Raghavan, S. V., & Liong, S.-Y. (2015). Statistical downscaling rainfall using artificial neural network: Significantly wetter Bangkok? *Theoretical and Applied Climatology*, *126*(3–4), 453–467. https://doi.org/10.1007/s00704-015-1580-1

Wang, N., Lombardo, L., Tonini, M., Cheng, W., Guo, L., & Xiong, J. (2021). Spatiotemporal clustering of flash floods in a changing climate (China, 1950–2015). *Natural Hazards and Earth System Sciences*, *21*(7), 2109–2124. https://doi.org/10.5194/nhess-21-2109-2021

Wilby, R. L., Wigley, T., Conway, D., Jones, P., Hewitson, B., Main, J., & Wilks, D. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, *34*(11), 2995–3008. https://doi.org/10.1029/98wr02577

Wu, J., & Gao, X. J. (2013). A gridded daily observation dataset over China region and comparison with the other datasets. *Chinese Journal of Geophysics*, *56*(4), 1102–1111. https://doi.org/10.6038/cjg20130406

Xiao, S., Xia, J., & Zou, L. (2020). Evaluation of multi-satellite precipitation products and their ability in capturing the characteristics of extreme climate events over the Yangtze River Basin, China. *Water*, *12*(4), 1179. https://doi.org/10.3390/w12041179

Yan, L., Lei, Q., Jiang, C., Yan, P., Ren, Z., Liu, B., & Liu, Z. (2022). Climate-informed monthly runoff prediction model using machine learning and feature importance analysis. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.1049840

Yan, L., Xiong, L., Jiang, C., Zhang, M., Wang, D., & Xu, C. Y. (2021). Updating intensity–duration–frequency curves for urban infrastructure design under a changing environment. *WIREs Water*, *8*(3). https://doi.org/10.1002/wat2.1519

Yang, L., Ma, J., Wang, X., & Tian, F. (2022). Hydroclimatology and hydrometeorology of flooding over the eastern Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, *127*(15), e2022JD037097. https://doi.org/10.1029/2022jd037097

Yao, T., Masson-Delmotte, V., Gao, J., Yu, W., Yang, X., Risi, C., et al. (2013). A review of climatic controls on $\delta^{18}O$ in precipitation over the Tibetan Plateau: Observations and simulations. *Reviews of Geophysics*, *51*(4), 525–548. https://doi.org/10.1002/rog.20023

Yin, S., & Chen, D. (2020). Weather generators. In *Oxford research encyclopedia of climate science* (pp. 1–45). https://doi.org/10.1093/acrefore/9780190228620.013.768

You, Q., Chen, D., Wu, F., Pepin, N., Cai, Z., Ahrens, B., et al. (2020). Elevation dependent warming over the Tibetan Plateau: Patterns, mechanisms and perspectives. *Earth-Science Reviews*, *210*, 103349. https://doi.org/10.1016/j.earscirev.2020.103349

Yuan, Y., Li, B., Gao, X., Liu, W., Li, Y., & Li, R. (2022). Validation of cloud-gap-filled snow cover of MODIS daily cloud-free snow cover products on the Qinghai–Tibetan Plateau. *Remote Sensing*, *14*(22), 5642. https://doi.org/10.3390/rs14225642

Zhang, G. (2019). Dataset of river basins map over the TP (2016) [Dataset]. National Tibetan Plateau Data Center. https://doi.org/10.11888/BaseGeography.tpe.249465.file

Zhang, G., Yao, T., Xie, H., Kang, S., & Lei, Y. (2013). Increased mass over the Tibetan Plateau: From lakes or glaciers? *Geophysical Research Letters*, *40*(10), 2125–2130. https://doi.org/10.1002/grl.50462

Zhang, J., Zhao, T., Li, Z., Li, C., Li, Z., Ying, K., et al. (2021). Evaluation of surface relative humidity in China from the CRA-40 and current reanalyses. *Advances in Atmospheric Sciences*, *38*(11), 1958–1976. https://doi.org/10.1007/s00376-021-0333-6

Zhang, Y. (2019). Integration dataset of Tibet Plateau boundary [Dataset]. National Tibetan Plateau Data Center. https://doi.org/10.11888/Geogra.tpdc.270099

Zhang, Y., Liu, L., Li, B., & Zheng, D. (2021). Comparison of boundary datasets covering Tibetan Plateau between 2021 and 2014 versions. *Journal of Global Change Data & Discovery*, *5*(3), 322–332. https://doi.org/10.3974/geodp.2021.03.10

Zhao, G., Liu, R., Yang, M., Tu, T., Ma, M., Hong, Y., & Wang, X. (2022). Large-scale flash flood warning in China using deep learning. *Journal of Hydrology*, *604*, 127222. https://doi.org/10.1016/j.jhydrol.2021.127222

Zhu, Y., Sang, Y.-F., Chen, D., Sivakumar, B., & Li, D. (2020). Effects of the South Asian summer monsoon anomaly on interannual variations in precipitation over the South-Central Tibetan Plateau. *Environmental Research Letters*, *15*(12), 124067. https://doi.org/10.1088/1748-9326/abc71b

Zipser, E. J., & Xu, W. (2011). Diurnal variations of precipitation, deep convection, and lightning over and east of the eastern Tibetan Plateau. *Journal of Climate*, *24*(2), 448–465. https://doi.org/10.1175/2010jcli3719.1