Earth System
Open Access
Science
Data

A dataset of standard precipitation index reconstructed from multi-proxies over Asia for the past 300 years

Yang Liu[1], Jingyun Zheng[1,2], Zhixin Hao[1,2], and Quansheng Ge[1,2]

[1]Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural
Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence: Jingyun Zheng (zhengjy@igsnrr.ac.cn) and Quansheng Ge (geqs@igsnrr.ac.cn)

Abstract. Proxy-based precipitation reconstruction is essential to study the inter-annual to decadal variability and underlying mechanisms beyond the instrumental period that is critically needed for climate modeling, prediction and attribution. Based on 2912 annually resolved proxy series mainly derived from tree rings and historical documents, we present a set of standard precipitation index (SPI) reconstructions for each year (November–October), covering the whole of Asia, and for the wet season (i.e., November–April for western Asia and May–October for the others) since 1700, with the spatial resolution of 2.5°. To screen the optimal candidate proxies for SPI reconstruction in each grid from available proxies in its connected region with a homogeneous rainfall regime and similar precipitation variability, a new approach is developed by adopting the grid-location-dependent division derived from the instrumental SPI data. The validation shows that these reconstructions are effective for most of Asia. The assessment of data quality compared with gauge precipitation before calibration time indicates that our reconstruction has high quality to show the precipitation variability in most of the study areas, except for a few grids in western Russia, the coastal area of southeast Asia and northern Japan. The full dataset can be obtained from https://doi.org/10.57760/sciencedb.01829 (Y. Liu et al., 2022).

1 Introduction

Asia bears the brunt of flood and drought disasters and is associated more than any other continent with extensive social and economic damages due to its large and heterogeneous landmass plus its high population densities in the southern and eastern regions (Lee et al., 2020; Wei et al., 2020). At the national level, 7 of the top 10 countries in the world with the largest number of population affected by climate-related disasters (mainly flood and drought) are located in Asia (CRED and UNISDR, 2015). However, the inter-annual, decadal and centennial spatiotemporal variability of Asian precipitation and the underlying mechanisms have not been fully characterized, which limits the performance of precipitation projection for the next decades to hundred years (Seth et al., 2019; Wang et al., 2021; F. Liu et al., 2022). Long-term, spatially resolved and high-quality precipitation datasets are needed to address these issues. Unfortunately, the global precipitation

observation network only covers the past century (Sun et al., 2018), while the data for the first half period in Asia is at low confidence levels (Hartmann et al., 2013). Therefore, proxy-based precipitation reconstructions are essential to quantify the precipitation variability beyond the instrumental period.

Up to now, there have been four gridded datasets to reconstruct summer (or the warm season) precipitation variability in mid–low latitude Asia for the past hundreds of years (Cook et al., 2010a; Feng et al., 2013; Shi et al., 2018, 2017) by using tree-ring chronologies only or by merging multi-proxies. For example, using 327 tree-ring chronologies mainly located in the Tibetan Plateau and Mongolia, Cook et al. (2010a) reconstructed the gridded (2.5° × 2.5°) summer (June–August, JJA) Palmer drought severity index (PDSI) over monsoon Asia during 1300–2005. By weighted merging 453 tree-ring-width chronologies and 71-site dryness–wetness grade series derived from Chinese historical docu-

ments (local gazettes), Shi et al. (2018) reconstructed a gridded Asian summer precipitation dataset for 1470–2013. Similar reconstructions were also conducted for North America (Cook et al., 2010b; Stahle et al., 2020), Europe (Cook et al., 2015, 2020) and for Oceania (Palmer et al., 2015). Moreover, by using a data assimilation (DA) approach to combine 2978 proxy data with the physical constraints of the atmosphere–ocean climate model together, a globally gridded ($2.0° \times 2.0°$) hydroclimate index dataset over the Common Era was also reconstructed (Steiger et al., 2018), including PDSI and the standardized precipitation evapotranspiration index (SPEI) for JJA, DJF (December–February) and April to the next March. These datasets extend records back in time and provide valuable efforts on improving the gridded paleoclimate reconstruction by synthesizing multiproxies from individual sites with spatiotemporal inhomogeneity.

However, intercomparisons of the abovementioned four gridded precipitation–drought variability reconstructions in monsoon Asia (Cook et al., 2010a; Feng et al., 2013; Shi et al., 2017, 2018) with independent instrumental observation data show notable differences among them, caused by proxies and methods for calibration, particularly dominated by the number and sample distribution of proxies used, as well as the seasonal sensitivity of the individual proxy to precipitation anomalies (Liu et al., 2021). For example, in the reconstruction from only tree-ring proxies, the explained variance in regions with sparse proxies (e.g., eastern China, mainland southeast Asia) is usually less than 20 % (Cook et al., 2010a). By merging tree-ring and documentary proxies in the reconstruction, the result is believed to illustrate large-scale rainfall variability faithfully but has more uncertainties in representing regional rainfall anomalies (Shi et al., 2018). Moreover, the precipitation over Asia has a complex spatial pattern, with the temporal variability on intra-seasonal and inter-annual scales (Hsu et al., 2014) due to different rainfall regimes in space (Awan et al., 2015; Conroy and Overpeck, 2011). Therefore, the sensitivity of individual proxies to precipitation anomalies has evident regional differences between seasons. In addition, many new proxies achieved in recent years are not utilized in the above-mentioned four gridded reconstructions in monsoon Asia (e.g., Shah et al., 2007; Sass-Klaassen et al., 2007; Arsalani et al., 2018, 2015; Chen et al., 2016; Zhang et al., 2017; Pumijumnong et al., 2020; Xu et al., 2015; Buckley et al., 2017; Ukhvatkina et al., 2021; Akkemik et al., 2020; Kostyakova et al., 2017; Kucherov, 2010; Xu et al., 2013; Borgaonkar et al., 2010). All of these motivate us to initiate this new gridded ($2.5° \times 2.5°$) reconstruction effort on seasonal to annual precipitation variability over the past 300 years in the whole of Asia, including the western and northern Asia not covered in the four gridded datasets developed in previous studies (Cook et al., 2010a; Feng et al., 2013; Shi et al., 2018, 2017). Noting that most of Russian territory is located in Asia, to keep the data integrity

at the national level, the whole Russian territory is included in this study.

2 Data and method

2.1 The study area and the framework for grid SPI reconstruction

The spatial coverage of our reconstruction is shown in Fig. 1, and the reconstructed target is standard precipitation index (SPI). In this vast study area, there are many climatic types with heterogeneous precipitation; specifically, the wet season in western Asia and the southwestern part of central Asia is mainly from November–April, but that in the rest of the regions is May–October due to different rainfall regimes in different regions (Bombardi et al., 2019; Peng et al., 2020). Thus, we reconstructed the annual (November–October) SPI for the entire study area, as well as the November–April SPI in western Asia and the southwestern part of central Asia and the May–October SPI in the other regions for the wet season. The flow chart of the reconstruction procedures is shown in Fig. 2. It is noted that there exists a complex spatial coherence pattern for the precipitation variation on scales of inter-annual, decadal and longer in the study area, which means the spatial representativeness of the individual proxy is dominated by the location in the context of the region (e.g., shape and area) with coherent rainfall regime and variation. Therefore, we develop a new approach to select proxies for each grid SPI reconstruction by adopting the grid-location-dependent division (GLDD) derived from the instrumental SPI data, instead of selecting proxies usually from an isotropic search radius for all grids in many previous studies (e.g., Cook et al., 2010a; Shi et al., 2018). The methods for reconstruction (including GLDD), searching candidate proxies for calibration, and validation will be presented in Sect. 2.4.

2.2 Instrumental data for calibration and spatial pattern of wet season identification

In our study, the grid size for SPI reconstruction is set as $2.5° \times 2.5°$. The instrumental data used for calibration are resized from the $0.5° \times 0.5°$ gridded monthly SPI data for 1948–2019 calculated by NOAA's land precipitation product (Chen et al., 2002), which was downloaded via IRI/LDEO Climate Data Library (http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/, last access: 21 May 2022). As pointed out by previous studies (Bombardi et al., 2019; Peng et al., 2020; Nieto et al., 2019), moisture sources are different across Asia throughout the year, and the wet season could be roughly classified as two terms of November–April and May–October. Therefore, to identify the spatial pattern of the wet season for SPI reconstruction in the $2.5° \times 2.5°$ gridded scale induced by different regional rainfall regimes, the monthly precipitation data for
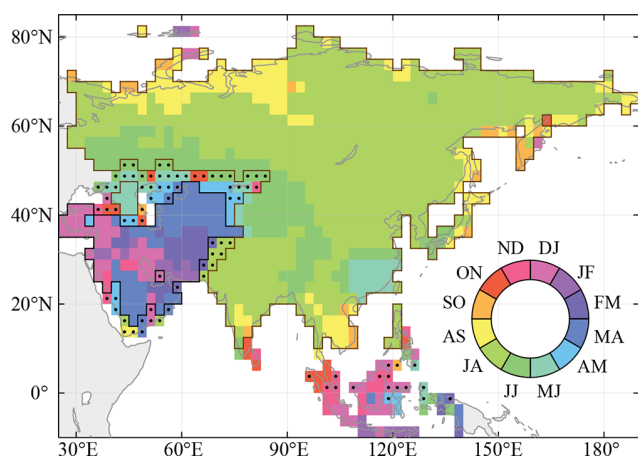
Figure 1. The study area and the spatial difference of rainfall regime, with the wettest bimester (two consecutive months) shown by the monthly GPCP precipitation data from 1948–2019. The dot marker indicates that the grid lacks a clear wet season. Annual (November–October) SPI is reconstructed for all non-gray grids, while wet season (November–April and May–October) SPI is reconstructed in regions with black and brown boundaries respectively.

1948–2019 by GPCC (Schneider et al., 2017) is also used to calculate two consecutive months with the most rainfall amount in a year (Fig. 1). It is shown that, in most parts of the study area, the wettest two consecutive months are in May–October. However, in western Asia (excluding the south corner of the Arabian Peninsula), the southwest part of central Asia and the tropical zone south to 10° N, the wettest two consecutive months are in November–April. Moreover, there also exist a few grids (dot marked in Fig. 1) that have no distinct wet season (Bombardi et al., 2019). Thus, we exclude the dotted grids in wet season SPI reconstruction.

2.3 Proxy data preparation

There are a total of 2912 annually resolved proxy series from Asia and adjacent land areas (Eastern Europe and Alaska) for reconstruction, of which 2792 are derived from tree-rings, 115 from historical documents, 4 from ice cores, and 1 from a stalagmite. Their spatial and temporal distribution is shown in Fig. 3. It is noted that all of the proxy series have at least 20 records overlapped with the instrumental period since 1948 to ensure a sufficient sample size for calibration and validation and more than 30 records before 1948 for reconstruction. The data source and standardized processes for each type of proxy series are described below.

Tree-ring data are mainly (2772) from the International Tree-Ring Data Bank (ITRDB), maintained by the World Data Center for Paleoclimatology (WDC-P, https://www.ncei.noaa.gov/products/paleoclimatology, last access: 28 October 2022), including 1854 tree-ring width records, 828 tree-ring density records, 67 tree-ring latewood percent

records, 22 tree-ring stable oxygen isotope (δ^{18}O) records and 1 tree-ring blue-intensity record. Most sites have two categories of data, i.e., original raw tree-ring measurements and tree-ring index chronologies derived from raw measurements. However, the index chronologies are not used directly in this study because they were standardized by various methods which are not described in the online metadata, and some of the methods may result in a substantial loss of long-term fluctuations (Coulthard et al., 2020). To maximally preserve the climatic-related low-frequency variance, we recalculate the chronologies from 2644 available raw measurement files by removing the growth trend with age-dependent splines (Melvin et al., 2007). In a few cases where age-dependent splines contain zeros or negative values, a more flexible curve, the Friedman variable span smoother (Friedman, 1984), is used to fit the growth trend. In addition, some trees experience disturbances during their lifespan, which could cause abrupt growth increases or reductions (Altman, 2020). To eliminate this effect, the running mean technique (Altman et al., 2014) is applied to identify the disturbance event, then separate growth curves are fitted before and after this year. Finally, the 51-year sliding expressed population signal (EPS) is calculated, and the threshold of 0.85 is used to determine the first reliable year of a chronology. The above procedures are also applied for sites with raw measurements only. The other 128 tree-ring records from ITRDB only have chronologies; EPS is not available, and thus we use the minimum sample size of 5 to determine the first reliable year. Besides ITRDB, 17 tree-ring width chronologies and 3 tree-ring δ^{18}O chronologies that indicate local precipitation or drought from recently published papers are included in our study (Shah et al., 2007; Sass-Klaassen et al., 2007; Arsalani et al., 2018, 2015; Chen et al., 2016; Zhang et al., 2017; Pumijumnong et al., 2020; Xu et al., 2015; Buckley et al., 2017; Ukhvatkina et al., 2021; Akkemik et al., 2020; Kostyakova et al., 2017; Kucherov, 2010; Xu et al., 2013; Borgaonkar et al., 2010). Compared with the tree-ring network used in previous studies over the monsoon Asia region (Cook et al., 2010a; Feng et al., 2013; Shi et al., 2017, 2018), a total of 113 ring-width chronologies are added in our study.

It is worth noting that part of the sites consists of both tree-ring width and density chronologies. According to the principle of dendroclimatology, the availability of soil water affects the growth rate and formation of wood, both within a season and the longer term; thus, tree-ring width is expected to be positively correlated with precipitation via this direct response (Vaganov et al., 2011; Wettstein et al., 2011). For tree-ring density chronologies, they are usually correlated with temperature variation and scarcely used in precipitation reconstruction (Briffa et al., 2002). However, due to multiple types of climate and complex topography in the vast study area, the tree-ring density chronologies and width chronologies with negative correlations to precipitation may also indicate precipitation variation well (George, 2014), and the use of such tree-ring predictors in hydro-climate reconstruction
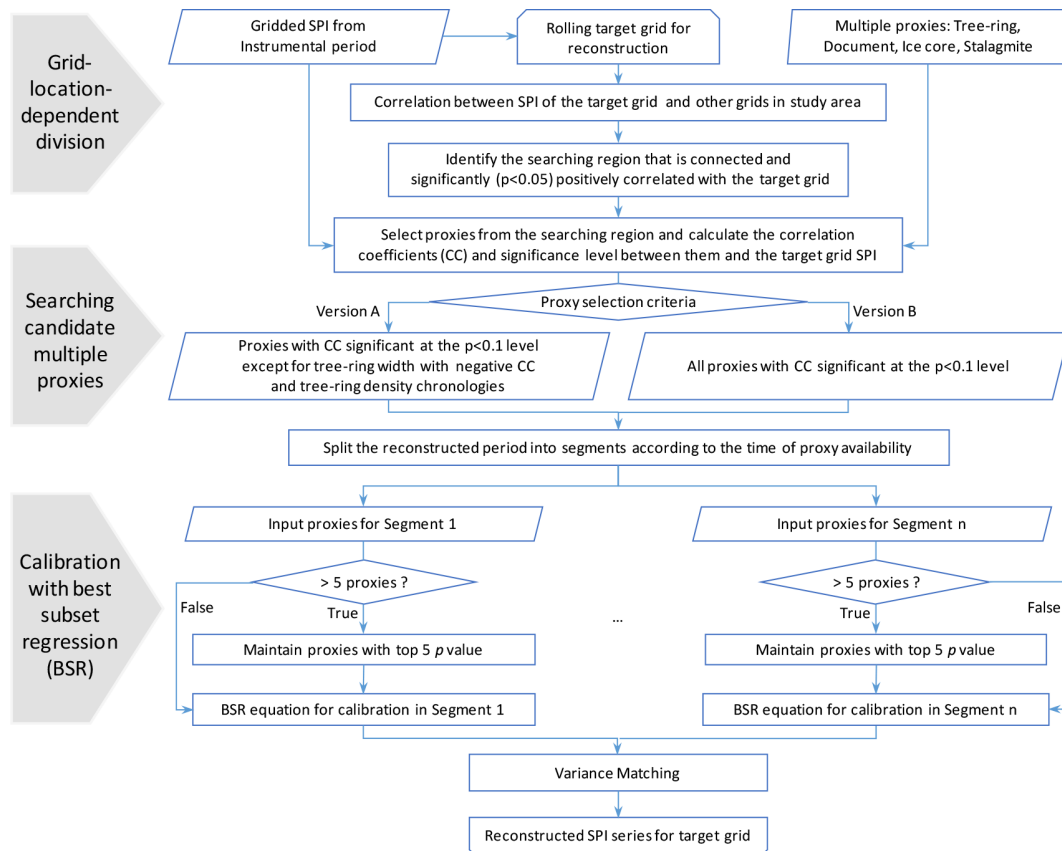
Figure 2. The flow chart of SPI reconstruction over Asia for 1700–2000 in this study.

has been discussed in prior studies (e.g., Cook et al., 2020). Therefore, we reconstruct two versions of SPI: one excludes tree-ring width chronologies negatively correlated to precipitation and tree-ring density chronologies (hereafter called "Version A"), the other includes all tree-ring chronologies (hereafter called "Version B").

The proxy from historical documents is mainly the dryness–wetness grade series for 120 sites in China for the past 500 years (henceforth referred to as DW120) by the Chinese Academy of Meteorological Science (CAMS, 1981). The grades were calibrated based on descriptions of droughts and floods and their impacts during the wet season, mainly recorded in Chinese local gazettes, using ideal frequency criteria of all time, roughly 10 % for grades 1 and 5 (heavy flood and severe drought), 20 %–30 % for grades 2 and 4 (flood and drought), and 30 %–40 % for grade 3 (normal). This grade dataset originally ended in 1979 (CAMS, 1981) and was extended to 2000 (Zhang et al., 2003; Zhang and Liu, 1993), which becomes an essential dataset to reconstruct summer precipitation over the Asia monsoon domain (Feng et al., 2013; Shi et al., 2017, 2018). However, DW120 contains a large proportion of missing data because there are only 26 040 grade records since 1700 (Fig. 4a), which limits the spatial and temporal coverage of data for gridded SPI re-

construction. Therefore, we update this dataset by two steps. The first is adding the missing data in DW120 from another dryness–wetness grade dataset for 63 sites in central eastern China (DW63) developed by Zhang (1996). It is noted that all sites of DW63 are included in DW120, and the grading criteria for DW63 are the same as DW120. Since DW63 was reconstructed from more abundant historical documents (such as the drought and flood descriptions recorded in the memoirs and archives of the Qing Dynasty), it had fewer missing records, with 100 % data availability after 1700. Therefore, all missing records of DW120 in central eastern China are added from DW63, which supplements 2045 grade records in total. The second step is interpolation from the isoline map of DW120 for individual years when most sites have available data (CAMS, 1981), which supplements 4121 grade records. Since DW63 was reconstructed by the same grading criteria as DW120, both the 2045 added grade records from DW63 and the 4121 added records from the yearly isoline map of DW120 match with the original available data. This updated DW120 finally contains 32 206 grade records since 1700, which is a 23.7 % increase compared with the original version (Fig. 4b). Unfortunately, no data are available before the 20th century for 10 sites in western China and one site in northeastern China (cross marked in Fig. 4b); thus, only 109
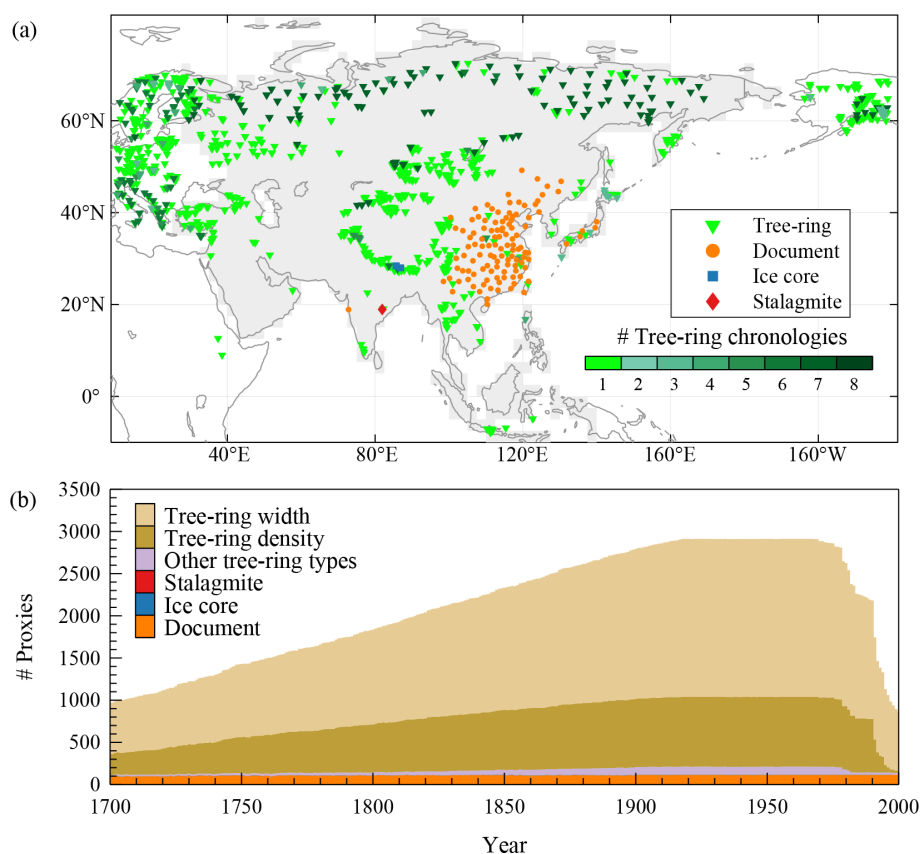
Figure 3. Spatial **(a)** and temporal **(b)** distribution of proxies.

sites in China are selected for our SPI reconstruction. Another documentary-based dryness–wetness grade series since 1781 is from Mumbai, India, which also consists of five grades calibrated against the percentage of rainfall anomalies derived from instrumental data in their overlapped period (Adamson and Nash, 2014). In addition, the series of wetseason (May–October) rainy days for five sites in Japan are also included. These series were extracted from the historical diaries (https://www.ncei.noaa.gov/access/paleo-search/study/5412, last access: 25 October 2022) and merged with instrumental data (Kamiguchi et al., 2010) by the method from Murata (1992).

The rest of the proxy series, derived from four ice cores in the Himalayas and one stalagmite in India, are also downloaded from WDC-P and have been proven to indicate hydroclimatic change by prior studies (Thompson et al., 2000; Sinha et al., 2011; Qin et al., 2002). It is worth noting that the $\delta^{18}O$ ratio series of the ice core from East Rongbuk Glacier is unequally spaced, with a mean temporal resolution of 0.082 years, which is simply re-sampled to an annually resolved series by averaging data in the same year.

2.4 Method for grid SPI reconstruction with grid-location-dependent division

Since there are many climatic types with heterogeneous precipitation in the study area, and since the spatial representativeness of an individual proxy is sensitive to location, we develop a new approach to identify the region for searching proxies (called "searching region" hereafter) to reconstruct SPI in each grid. This approach is developed according to the regional division of the coherence of inter-annual precipitation variations in the context of the spatial pattern of rainfall regimes to ensure the proxies in the searching region can indicate SPI variability in the target grid well. We divided the regions from the spatial pattern of the correlation coefficient (CC) between the SPI of each target grid for SPI reconstruction and the other grids within the study area, which is calculated from instrumental SPI data. The searching region is defined as all connected grids surrounding the target grid with CCs passing 0.05 significance level. Thus, this searching region has the robust coherence of precipitation variability and rainfall regimes with the target grid, and the proxies in this region have the best spatial representativeness in relation to the target grid. Since this regional division is dependent on the grid location by rolling the target grid, this approach was called "grid-location-dependent division
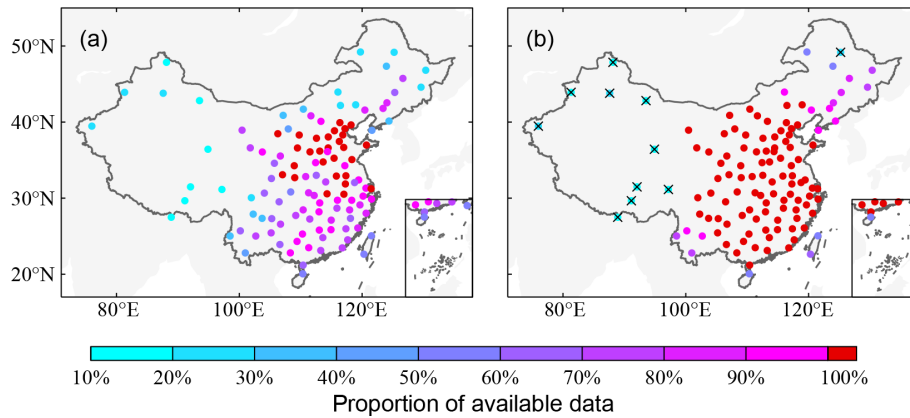
Figure 4. The proportion of available data for DW120 in the original version **(a)** and after updating **(b)**. Sites with a cross marker in **(b)** are excluded in the reconstruction.

(GLDD)" in our study (Fig. 2). Moreover, we use best subset regression (BSR) to identify optimal combinations of the candidate proxies for calibration in each grid SPI reconstruction based on available proxies in different intervals (Fig. 2) because the proxies are unevenly distributed in space and time. The following shows the November–October SPI reconstruction for a grid of 90.0–92.5° E and 27.5–30.0° N (located in southwest China) as an example of the detailed steps.

Firstly, the spatial SPI correlation field of the target grid is calculated, and the regions with positively significant ($p < 0.05$) correlation coefficients are identified (Fig. 5a). It shows that the target grid and its adjacent grids have significant correlations that cover an irregular shape (i.e., not a circle-like shape, with an isotropic radius from the target grid, or other regular shape) in southwest China, which means there exists robust coherence for SPI variation. This is because the rainfall regime and precipitation for that region are usually dominated by the same atmospheric circulation systems (Zhang and Wang, 2021). Besides, there are some other remote regions (e.g., the Malay Archipelago, the Russian Plain, and regions around the New Siberian Islands) that show significant correlations to the target grid. However, prior studies have reported that the long-distance precipitation teleconnection patterns are usually unstable over a long-term period, since they are linked by large-scale atmospheric circulations or propagating waves (Wu, 2016; Boers et al., 2019). Therefore, the candidate proxies for the target grid SPI reconstruction should be searched only from the connected region (Fig. 5b), and there are a total of 43 proxies for the candidate proxy selection.

Secondly, the correlations between the target grid SPI and each series of 43 proxies in the searching region are calculated to select the candidate proxies by the threshold of the 0.1 significance level for the correlations. Note that prior summer precipitation could affect the tree-ring formation in the next year (Wettstein et al., 2011); thus 1-year-lagged tree-ring chronologies are also included for the

November–October SPI reconstruction. However, the proxies with highly positive correlations may lead to multi-linearity effects in the regression equation for calibration. Thus, we also calculate the correlations among all 43 proxy series, and if any pair of proxy series shows an extremely high positive correlation (i.e., $r > 0.90$ and $p < 0.0001$) in their common period, the shorter one will be excluded from the pool of candidate proxies. By this step, a total of eight proxy series (including five tree-ring width series, one tree-ring δ^{18}O series and two dryness–wetness grade series) are selected for BSR in the following step (Fig. 5b).

Thirdly, the calibration equation is established by using BSR for each time segment, depending on the length of the candidate proxy series. According to the start and end year of all eight candidate proxy series, the time of proxy availability should be classified into six segments, in which there are eight candidate proxies for 1772–1997, seven candidate proxies for two segments in 1745–1771 and 1998–2000 respectively, six candidate proxies for 1743–1744, five candidate proxies for 1739–1742 and four candidate proxies for 1700–1738 (Fig. 5c). Moreover, to avoid the overfitting in the regression induced by redundant independent variables (Lever et al., 2016), if there are more than five candidate proxies, only five proxy series with the top five significance levels for the correlations with target SPI are retained for the regression. This is because the sample length to develop calibration equations for reconstruction is about 50 years usually, and the sample size should preferably be 10 times (or more) the number of variables for BSR, according to the principle of statistics (Sekaran, 2003). Thus, three individual segments (1743–1744, 1745–1771 and 1772–1997) retain the same five proxies (i.e., two tree-ring width series, one tree-ring δ^{18}O series and two DW120 series), and they could be regarded as one segment of 1743–1997. Then, we use BSR to establish four calibration equations for SPI reconstruction in 1700–1738, 1739–1742, 1743–1997 and 1998–2000 respectively, in which the best subset selection is de-
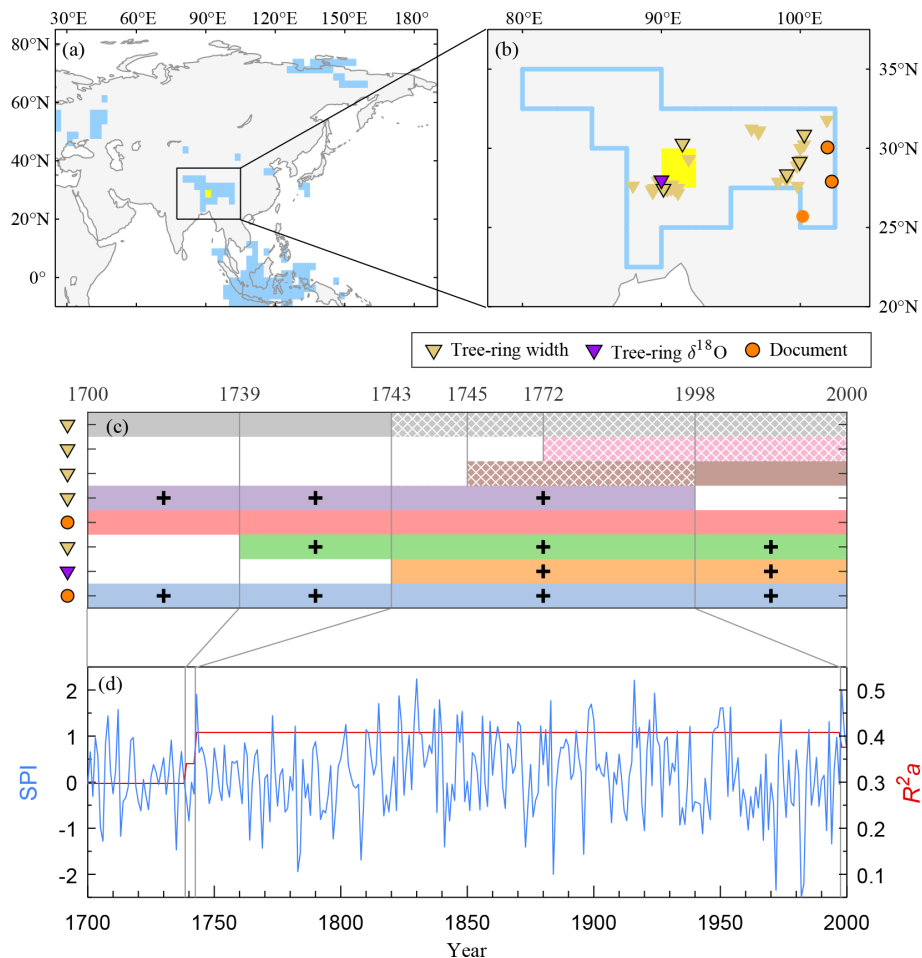
Figure 5. Demonstration of a grid SPI reconstruction for showing proxy selection by the GLDD approach. **(a)** The target grid (yellow square) and regions (light blue) that have significantly (at least $p < 0.05$) positively correlated SPI change. **(b)** The searching region connected with the target grid and proxies in it. A proxy marker with a black edge means it is significantly (at least $p < 0.1$) correlated with SPI change in the target grid. **(c)** Temporal coverage of picked proxy series and derived four segments based on available proxies. Proxies are listed in ascending order of p value from bottom to top. When a segment has more than five proxies, the bottom five (solid patch) are used in BSR, and the others (cross patch) are excluded. Proxies that remained in the final BSR model are marked with plus signs. **(d)** Reconstructed SPI series and calibration R^2a for each segment.

termined by maximizing the coefficient of efficiency (CE) (Cook et al., 1994), calculated by a state-of-the-art 4-fold rolling-window cross-validation procedure (Nguyen et al., 2020). Another commonly used validation parameter, reduction of error (RE), is also calculated from the same procedure. Finally, the target SPI series for the full time is constructed by merging the reconstructions for individual segments (Fig. 5d). As the reconstructions for different segments were calibrated from different equations with different variances and predicted sums of squares, the magnitudes of the reconstructed SPI for a specific segment had to be adjusted using the variance matching method with respect to the standard deviations of the predictands in common years during the calibration period (McCarroll et al., 2015).

3 Results and discussion

The dataset includes four SPI reconstructions: (1) the November–October SPI reconstruction for the whole of Asia without using tree-ring density chronologies and width chronologies with negative correlations to precipitation (November–October SPI Version A); (2) the November–October SPI reconstruction for the whole of Asia, adding tree-ring density chronologies and width chronologies with negative correlations to precipitation (November–October SPI Version B); (3) the wet-season SPI reconstruction for the extra-tropical Asia (November–April SPI for western Asia and May–October SPI for the rest of the regions) without using tree-ring density chronologies and width chronologies with negative correlations to precipitation (wet season SPI Version A); (4) the wet season SPI reconstruction for the

extra-tropical Asia (November–April SPI for western Asia and May–October SPI for the rest of the regions), adding tree-ring density chronologies and width chronologies with negative correlations to precipitation (wet season SPI Version B). Each of them is stored in a NetCDF file (.nc) and contains five three-dimensional (longitude × latitude × time) variables, including reconstructed SPI, adjusted coefficient of determination (R^2a), validation RE, validation CE and the number of proxies used for construction (nPrx).

3.1 Validity of the reconstruction

Figure 6 show the spatial patterns of R^2a, RE and CE for the November–October SPI reconstruction since 1700 by a 50-year interval. It shows that CE in most of the study areas is positive. Although a few grids have negative CE, especially before 1800, most of them still have positive RE. These results mean the reconstruction is effective, in which the area with $R^2a > 0.2$ accounts for 36.1 % of grids in Asia in 1700 and extends to 66.1 % in 1950 due to more and more available proxies. Since 1700, the areas with $R^2a > 0.4$ are distributed in a broad region from the southwest coast of the Caspian Sea to Balkhash Lake to eastern China and in some grids in the northern Far East, northern India and the western mainland Southeast Asia. R^2a gradually passed 0.4 from 1750 to 1800 over Turkey, the West Siberian Plain, central Asia, Mongolia and India. The highest R^2a (more than 0.6) appeared in central eastern China throughout the entire 300-year period.

By adding tree-ring density chronologies and width chronologies with negative correlations to precipitation in the November–October SPI reconstruction (i.e., Version B), the number of grids with ineffective reconstruction is significantly reduced, and the R^2a for most of the grids is significantly increased (Fig. 7). Compared to the November–October SPI reconstruction Version A (Fig. 6), the area with $R^2a > 0.2$ in Version B accounts for 59.3 % of grids in Asia in 1700 and extends to 85.1 % in 1950. In particular, R^2a increased by 0.2–0.3 in central to eastern Russia and by 0.1 to 0.2 in other regions, except for the Arabian Peninsula and eastern China.

Likewise, for the wet-season SPI reconstruction, it is also effective in most grids (Fig. 8), in which the area with $R^2a > 0.2$ accounts for 37.3 % of grids in Asia in 1700 and extends to 61.5 % in 1950. Compared with the November–October SPI Version A (Fig. 6), the wet-season SPI reconstruction shows significantly higher R^2a (0.1–0.2) for the region on the east of the Caspian Sea, slightly higher R^2a (around 0.1) for most grids in high-latitude zones, and a reduced R^2a around 0.1 in eastern China. For the wet-season SPI reconstruction with added tree-ring density chronologies and width chronologies with negative correlations to precipitation, the percentage of areas with $R^2a > 0.2$ in 1700 and 1950 is 58.1 % and 83.9 % respectively (Fig. 9). The difference in skill metrics between two wet-season SPI versions

(Figs. 8 and 9) is similar to that between two November–October SPI versions (Figs. 6 and 7).

3.2 Data quality and usability

Compared to three reconstructions of summer (JJA or May–September) precipitation (or PDSI) in monsoon Asia by previous studies (Cook et al., 2010a; Feng et al., 2013; Shi et al., 2018), the R^2a in the calibration period of our May–October SPI reconstructions (Figs. 8p and 9p) are 10 % higher than that of the best one in three reconstructions over the south Tibetan Plateau to the eastern India subcontinent, the western mainland Southeast Asia and northwest China. Moreover, our reconstruction has a slightly higher R^2a in parts of Mongolia, central Asia and eastern China than that in other reconstructions. In particular, R^2a in eastern China in our reconstruction is about 40 % higher than that from the reconstruction by only tree-ring data (Cook et al., 2010a). These improvements are not only because more proxy data (including the DWI derived from Chinese historical documents and the tree-ring data published recently) are added but also because of the development of the reconstruction method that selects proxies by the GLDD approach from a connected searching region with significantly positive correlations to the target grid SPI.

In addition, the maps of correlation between our wet-season SPI reconstructions and four reconstructions in monsoon Asia by previous studies show that most grids pass the 0.01 significance level (Figs. 10–11). Specifically, for the correlation (Figs. 10a, 11a) between our wet-season SPI reconstruction Versions A and B and the JJA precipitation reconstruction by Shi et al. (2018), 63.2 % and 64.1 % of all grids passed the 0.01 significance level, in which the value of the correlation coefficients for central eastern China are almost higher than 0.60. Similar results are also found for the correlations between our reconstruction versus the May–September precipitation anomaly reconstruction by Shi et al. (2017) in China (Figs. 10b, 11b) and the May–September precipitation reconstruction over monsoon Asia (Figs. 10c, 11c) by Feng et al. (2013). Even for the correlations between our wet-season SPI reconstruction versus JJA PDSI reconstruction for monsoon Asian (Figs. 10d, 11d) by Cook et al. (2010a) only using tree rings, 57.4 % (for our reconstruction Version A versus JJA PDSI reconstruction) and 58.8 % (for our reconstruction Version B versus JJA PDSI reconstruction) of all grids passed the 0.01 significance level.

To further assess the quality of reconstructed data, we compare our November–October SPI reconstruction with gauge precipitation at those weather stations with at least 30-year records before 1948, in which the precipitation data are from the Global Historical Climatology Network monthly dataset version 2 (GHCNmv2, https://www.ncei.noaa.gov/pub/data/ghcn/v2/, last access: 21 May 2022) and the long-term instrumental climatic databases of the People's Republic of China (Tao et al., 1997). We calculate the correlations
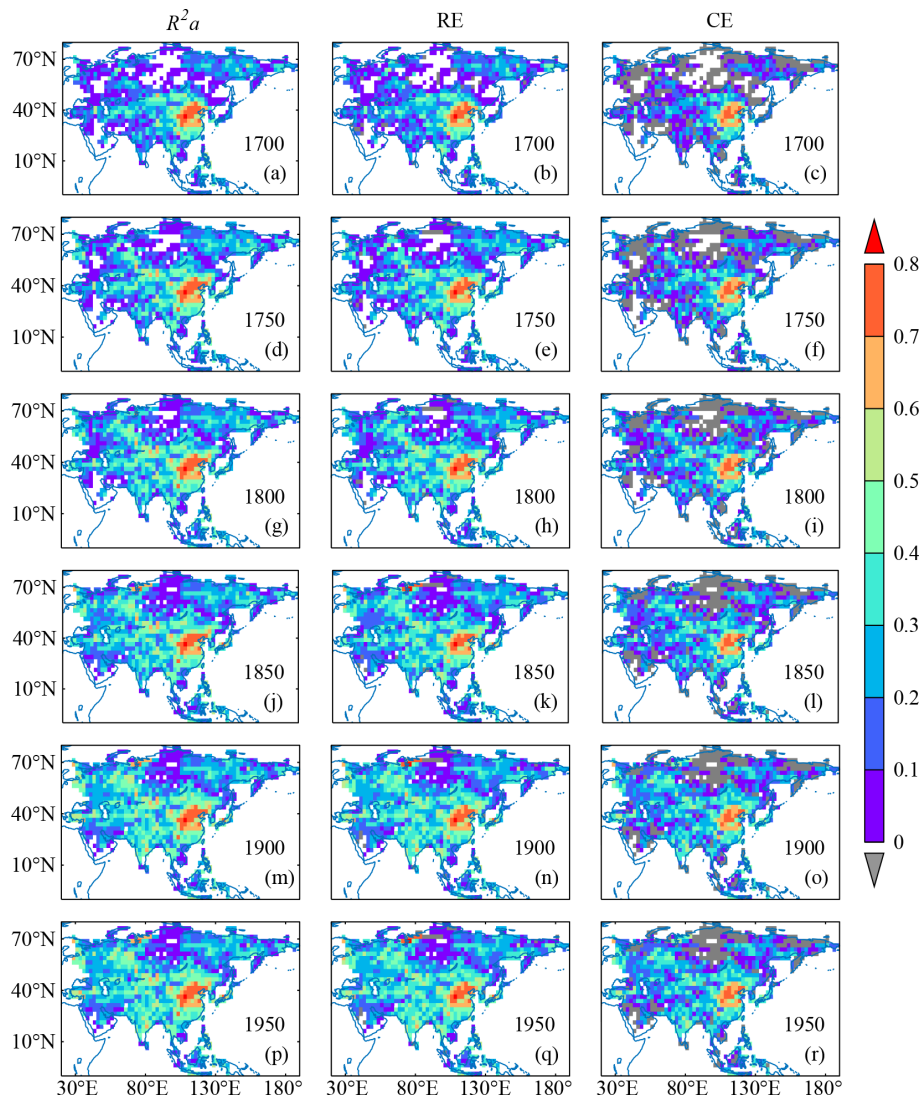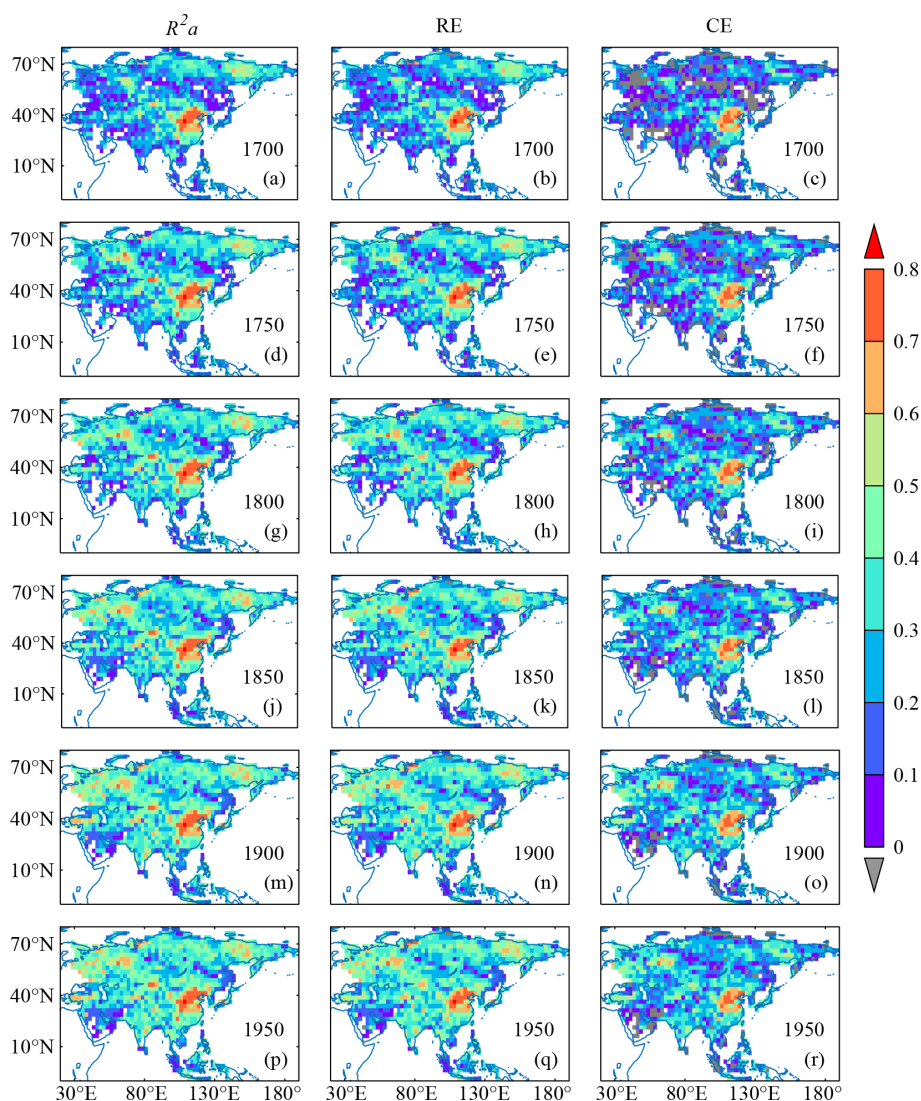
Figure 6. R^2a, RE and CE for November–October SPI reconstruction by multi-proxies without using tree-ring density chronologies and width chronologies with negative correlations to precipitation.

between the November–October precipitation anomaly percentage and November–October SPI reconstruction in corresponding grids (Fig. 12). Noting that the length of instrumental data before 1948 varies for different weather stations, i.e., the degrees of freedom for calculating these correlations are different station by station, we show the significance level for all positive correlations station by station instead of the value of correlation coefficients directly, with levels of $p \leq 0.01$, $0.01 < p \leq 0.05$, $0.05 < p \leq 0.1$ and $p > 0.1$ (Fig. 12). The result shows that the correlations for most sites, especially in eastern and southern Asia, pass the significance level of 0.1, though the correlation is not significant for parts of stations in central Asia, western Asia, the coastal area in southeast Asia and western Russia. For example, in the six sites (Haerbin, Beijing, Qingdao, Shanghai, Yichang and Shantou) evenly distributed across eastern China (Fig. 12), all the

correlations between reconstruction and observation pass the 0.01 significance level (Fig. 13). Moreover, the reconstructions could reproduce the most extreme years, e.g., 1853, 1871, 1890, 1893, 1920 and 1921 in Beijing; 1875, 1876, 1889, 1891, 1892, 1921, 1929, 1931 and 1934 in Shanghai; and 1889, 1897, 1900, 1902, 1920, 1928, 1935 and 1937 in Yichang (Fig. 13). The high p values (i.e., $p > 0.1$) of the poor positive correlation or negative correlation in parts of stations (e.g., in India or southeast Asia) might be induced by uncertainties from reconstructions based on few available proxies and observations in early times because the instrumental data from these station usually extend to the 1880s and before (e.g., several stations in India extend to 1836), with missing records and the frequent use of defective rain gauges in the early times. This assessment indicates that our reconstruction is of high quality for showing the precipita-

https://doi.org/10.5194/essd-14-5717-2022

Earth Syst. Sci. Data, 14, 5717–5735, 2022

Figure 7. R^2a, RE and CE for November–October SPI reconstruction by multi-proxies, including tree-ring density chronologies and width chronologies with negative correlations to precipitation.

tion variability in most of the study areas, except for a few grids in western Russia, the coastal area of southeast Asia and northern Japan. Thus, these datasets could be used to further study the spatiotemporal variability and underlying mechanisms of Asian precipitation since the pre-industrial era that is critically needed for climate modeling, prediction and attribution.

For example, we use the dataset of November–October SPI Version B to investigate the spatiotemporal pattern of hydroclimate variability over Asia for 1700–2000 by empirical orthogonal function (EOF) analysis. Figure 14a shows the first 10 eigenvalues and their 95 % confidence uncertainty intervals generated by the method from North et al. (1982). We find that only the first leading eigenvalue is independent, while the uncertainty intervals of other eigenvalues are overlapped, and the cumulative explained variance of the first 10

eigenvalues only accounts for 30.83 % of the total (Fig. 14b). Such results indicate that there exist multiple spatial patterns of precipitation in Asia. Here, we show some major characteristics for the first four modes, including spatial patterns (Fig. 14c–f), temporal changes (Fig. 15) and their correlations with winter sea surface temperature anomalies (SSTAs) after the high-pass filter (Fig. 16). It is noted that the gridded SSTA data are from ERSSTv5 over 1854–2000 (Huang et al., 2017).

The spatial pattern of EOF1 has strong negative loadings over central Russia and a broad region from western Asia to central Asia to western China, while dominant positive loadings can be observed over the monsoon region and eastern Russia (Fig. 14c). The time series for EOF1 has powerful inter-annual fluctuations over the considered period and significant decadal fluctuations in 1770–1820 (Fig. 15a). Its
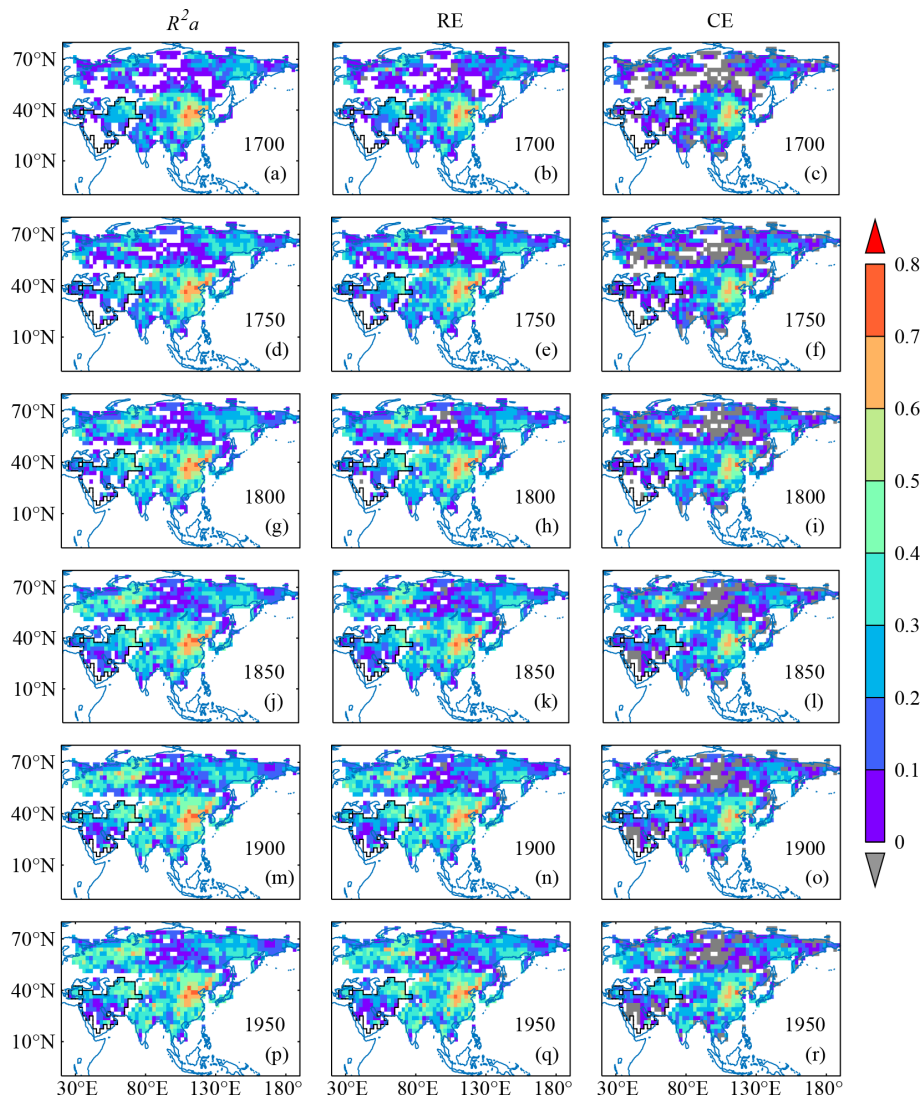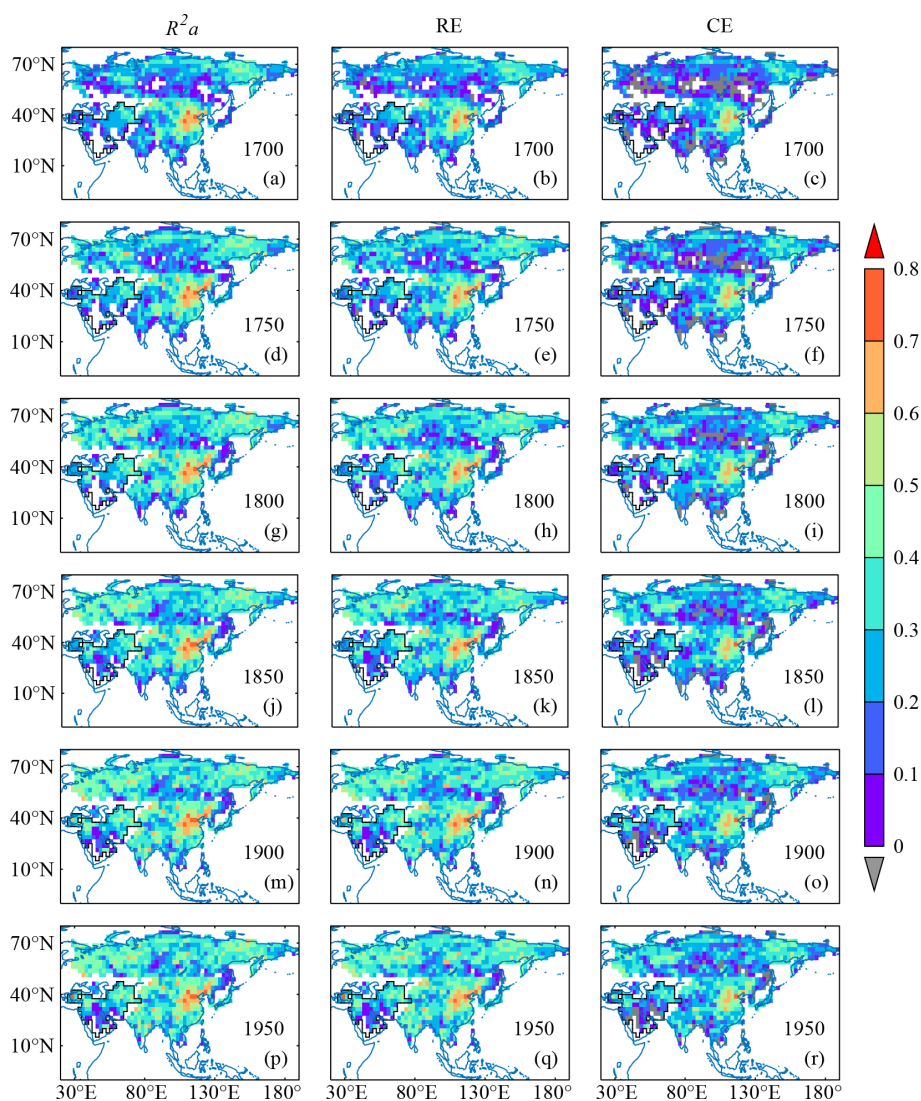
Figure 8. R^2a, RE and CE for wet-season SPI reconstruction by multi-proxies without using tree-ring density chronologies and width chronologies with negative correlations to precipitation; the black line indicates the boundary of the region in which the wet season is November–April, as that in Fig. 1.

high-frequency change (10-year high-pass filter) shows striking negative correlations with winter SSTA in the central equatorial Pacific and Indian Oceans but positive correlations in the western tropical Pacific Ocean (Fig. 16a), which suggests that this mode is strongly affected by coupling oscillation in tropical oceans, i.e., Indo-Pacific tripole (Lian et al., 2013).

The EOF2 shows a teleconnected pattern, with positive loadings over western Asia, India, northern China and western and eastern Russia but negative loadings in the rest of the regions (Fig. 14d). The energy bands of the time series for EOF2 are similar to those for EOF1, while its decadal fluctuations are expressed in 1840–1900 (Fig. 15b). The interannual fluctuation of this mode is significantly correlated with winter SSTA only in the eastern tropical Pacific Ocean

(Fig. 16b), which indicates that EOF2 is dominated by El Niño–Southern Oscillation (ENSO).

The EOF3 and EOF4 both show multi-pole spatial patterns (Fig. 14e–f), and their time series are dominated by decadal to multi-decadal scale fluctuations (Fig. 15c–d). They express some significant decadal precipitation patterns in specific regions. For instance, precipitation over eastern China has two major patterns of decadal variation (Zheng et al., 2016): one is a dipole pattern divided by the Huai River, and it is consistent with EOF3 in our study (Fig. 14e); the other is a four-zone pattern (centered in southern China, the Yangtze River Valley, the North China Plain and northeastern China) which is similar to EOF4 (Fig. 14f). In contrast, the high-frequency fluctuations are relatively weak for these two EOFs, and their links with SSTA are not significant.

Figure 9. R^2a, RE and CE for wet-season SPI reconstruction by multi-proxies, including tree-ring density chronologies and width chronologies with negative correlations to precipitation; the black line indicates the boundary of the region in which the wet season is November–April, as that in Fig. 1.

3.3 Effectiveness of GLDD and Uncertainty

Limited by the spatial coverage and uneven distribution of available proxies, to reconstruct the grid dataset on past climate for large scales such as continental, hemispherical or global, it is necessary to search the proxy for calibration from a large area (so-called "searching region", usually). In previous studies on past temperature (Christiansen and Ljungqvist, 2017) or hydro-climate reconstruction (Cook et al., 2010a; Shi et al. 2017, 2018), the searching region for each grid is set as a circular area with the same isotropic searching radius (ISR). However, as pointed out by Christiansen and Ljungqvist (2017) from the investigation on spatial decorrelation length in the Northern Hemispheric temperature field, the searching radius for different target grids

varies from less than 1000 to more than 6000 km, relying on target grid location and the searching direction along with the spatial pattern of coherence of temperature variation at different time scales. Since the spatial heterogeneity of precipitation variation is more evident than that of temperature variation, the proxy for a target hydroclimate reconstruction should be more sensitive to location and the searching direction.

Compared with searching proxies using isotropic searching radius, GLDD searches proxies for each target grid from the surrounding region where the precipitation variability and rainfall regime are robustly coherent with that of the target grid. As shown in Fig. 17, there are evident differences in the maximum (Fig. 17a) and minimum distances (Fig. 17b) from the boundary of the searching region to each target grid

Earth Syst. Sci. Data, 14, 5717–5735, 2022

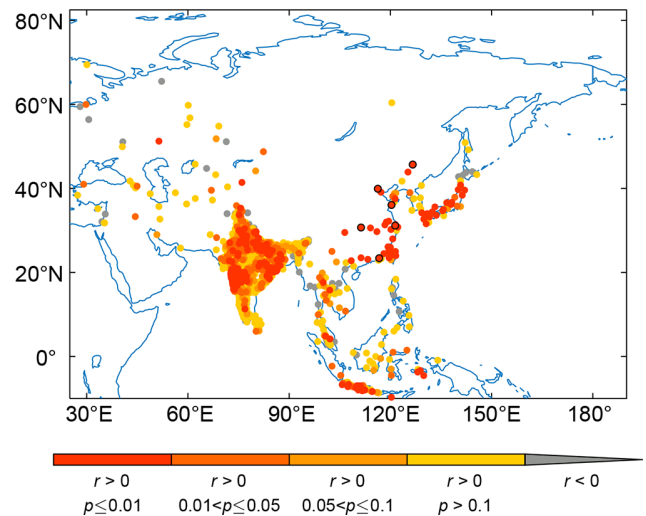https://doi.org/10.5194/essd-14-5717-2022
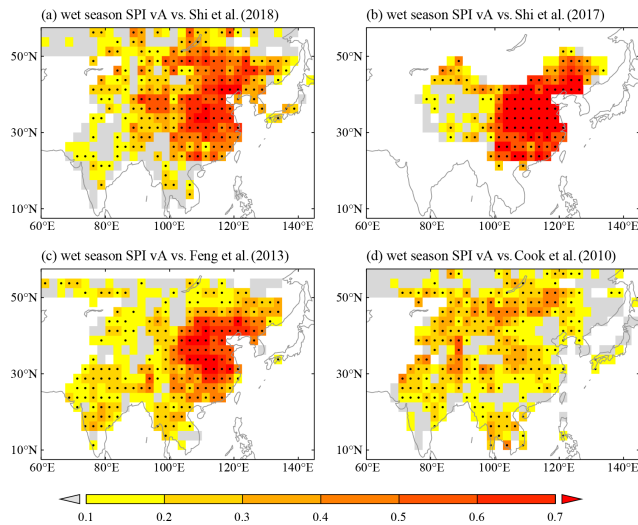
Figure 10. The maps of correlation between the wet-season SPI reconstruction Version A of this study and four reconstructions in monsoon Asia by previous studies of Shi et al. (2018) **(a)**, Shi et al. (2017) **(b)**, Feng et al. (2013) **(c)** and Cook et al. (2010a) **(d)** respectively. Correlation values significant at 99 % confidence are shown by dot marker.
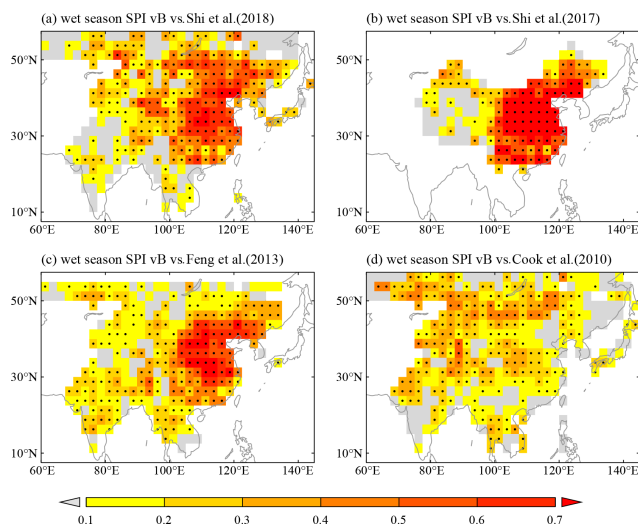

**Figure 11.** Same as Fig. 10 but for the wet-season SPI reconstruction Version B.

Figure 12. Correlations between November–October precipitation anomaly percentage for weather stations with at least 30-year records before 1948 from GHCNm and November–October SPI reconstruction in corresponding grids. The six selected sites in eastern China are shown with black edges, and the comparisons between observation and reconstruction year by year in these sites will be shown as examples in Fig. 13.
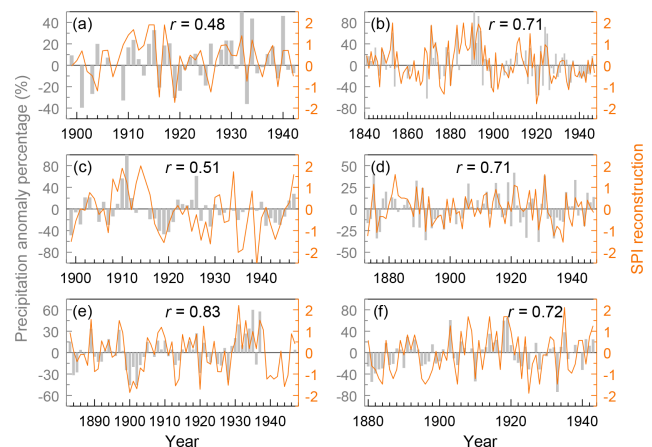

Figure 13. Comparisons between November–October precipitation anomaly percentage for six sites across eastern China from Tao et al. (1997) and November–October SPI reconstruction in corresponding grids. **(a)** Harbin (126.62° E, 45.68° N), **(b)** Beijing (116.28° E, 39.93° N), **(c)** Qingdao (120.33° E, 36.07° N), **(d)** Shanghai (121.43° E, 31.17° N), **(e)** Yichang (111.30° E, 30.70° N) and **(f)** Shantou (116.68° E, 23.40° N). Their locations are also shown in Fig. 12.

across Asia. The maximum distance is 1000–2000 km for most grids in China, Mongolia, and central and northwestern Russia; 2000–3000 km for most grids in India, central Asia and southwestern and eastern Russia; 3000–4000 km in the Arabian Peninsula; and more than 4000 km for tropical islands. However, the minimum distance is only 250–750 km for most grids of the study area, except very few grids in the tropics. The difference between the maximum and minimum distances (Fig. 17c) could reach 2000 km or more in regions with high topographic complexity, which means that

the searching region is always in an irregular shape. Thus, for the area (e.g., the Tibetan Plateau and surrounding area) with complicated topography and multiplex hydroclimate variation, GLDD could identify the unique searching region (including shape and size) rigorously for each target grid. For the area with a homogeneous hydroclimate variability and
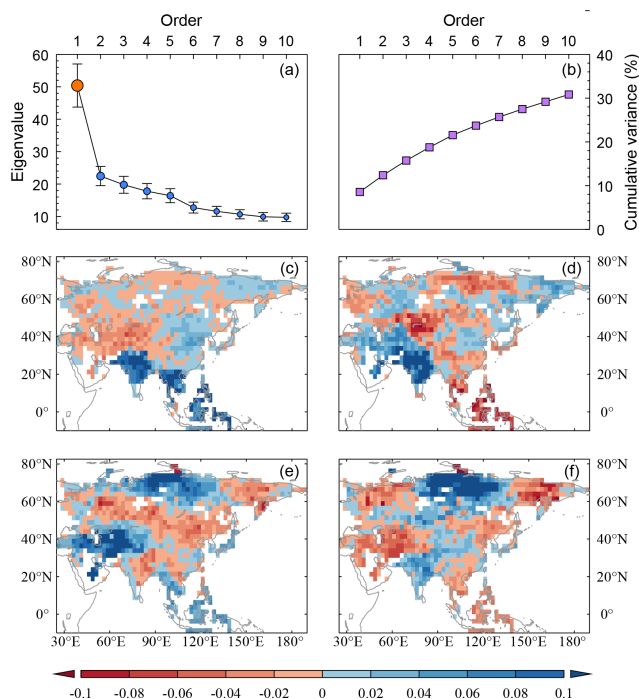
Figure 14. EOF analysis of November–October SPI reconstruction in Asia. **(a)** The first 10 eigenvalues and their 95 % uncertainty intervals. **(b)** The cumulative explained variance of the first 10 eigenvalues. **(c–f)** Spatial patterns of EOF1–EOF4.
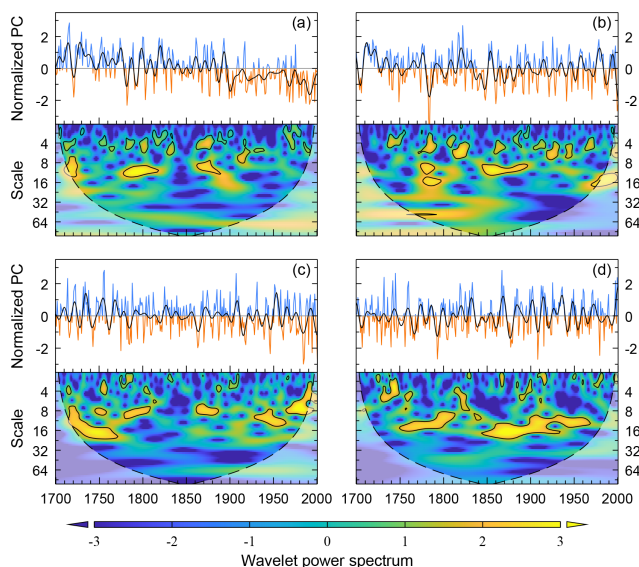

Figure 15. Temporal change and wavelet power spectrum of the time series (PC) for EOF1–EOF4 **(a–d)** shown in Fig. 14. PC is shown after normalization with a 10-year low-pass filter (black) applied to each. Spectral bands significant above the 90 % level are shown by black contours.

rainfall regime, GLDD could capture the proxies far from the target grid, which could reconstruct well in the areas where proxy data are not present, such as the east of the Caspian Sea. Therefore, GLDD could search the optimal proxies for hydroclimate reconstruction for each grid and consequently improve the quality of the reconstructed dataset (see the overall improvement of R^2a compared with prior studies in Sect. 3.2). For example, the R^2a of our reconstruction for the grid of 90.0–92.5° E and 27.5–30.0° N in Fig. 5 reached 40.8 %, which is 10.9 % higher than that of Shi et al.'s (2018) reconstruction by merging tree ring and documentary records together via ISR.

Yet, in this study, there still exist limitations and uncertainties. First, GLDD could only search the candidate proxies for the reconstruction in a target grid from its connected grids, where the precipitation variability and rainfall regime are robustly coherent with the target grid, and excludes the proxies from a teleconnected pattern. Thus, the candidate proxies for the target grids are usually limited by GLDD compared to the ISR approach. This might not only enlarge the uncertainty but also induce missing data in the reconstruction for an area with large spatial heterogeneity of precipitation variation and rainfall regimes due to very few available candidate proxies, such as central Russia and the Arabian Peninsula. Second, for tree-ring proxies, we use the same standardization method to build chronologies when raw measurements are available. However, about 4.5 % of the tree-ring proxies do not have raw measurement files; thus, we have to use the processed chronologies with various standardization methods from different data providers. As the test for some sites, the difference between chronologies could reach a maximum of 20 % from different standardization methods (Li et al., 2011). This may also induce uncertainty in the reconstruction. Thirdly, for documentary proxies, DW120 may use instrumental precipitation data to identify the dryness–wetness grades since 1951, especially after 1979, which might lead to overestimations of the calibration and verification metrics in eastern China. Fortunately, the data of dryness–wetness grades before 1950 are completely derived from historical documents (Wang and Zhao, 1979). Thus, comparisons between Figs. 12–13 and 6–7 by each site grid could help us to assess the overestimation, and the result shows that the overestimation of R^2a in this reconstruction is about 10 % on average over eastern China.

4   Data availability

The dataset (Y. Liu et al., 2022) can be accessed at https://doi.org/10.57760/sciencedb.01829. This dataset is licensed under a CC BY-SA 4.0 license.
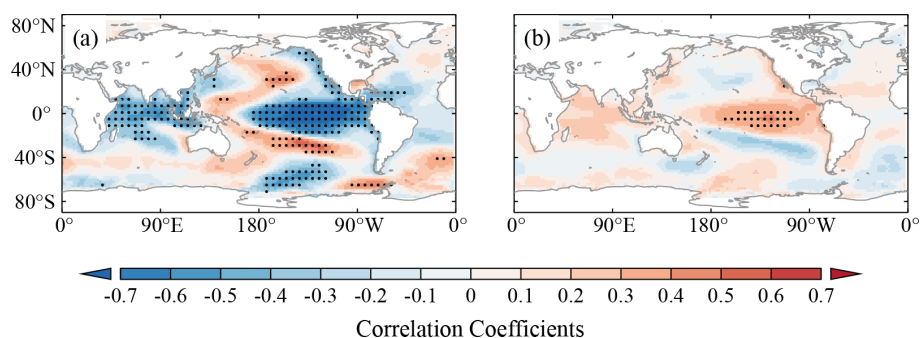
Figure 16. Field correlations between SSTA in winter and time series of EOF1 **(a)** and EOF2 **(b)** after the 10-year high-pass filter. Correlation values significant at 95 % confidence are shown by a dot marker.
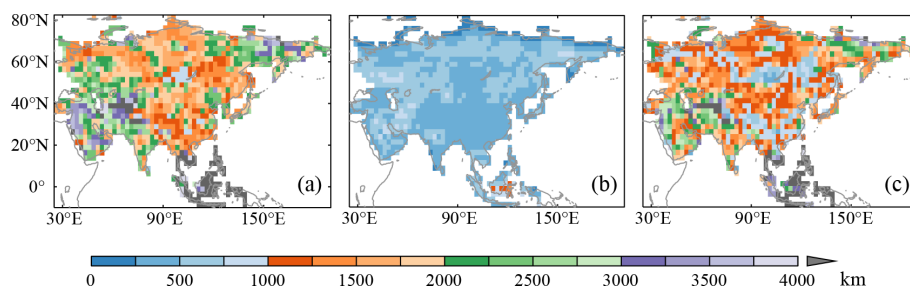
Figure 17. The maximum **(a)** and minimum **(b)** distance from the boundary of the searching region to the target point and their difference **(c)** for each grid.

5 Conclusions

In this study, we use a multi-proxy (mainly from tree rings and historical documents with clear annual dating) network containing 2912 series to reconstruct SPI for the wet season (November–April for west Asia and May–October for the other regions) and annual (November–October) timescales since 1700 over Asia, with a spatial resolution of $2.5° \times 2.5°$. Compared to the previous studies (Cook et al., 2010a; Feng et al., 2013; Shi et al. 2017, 2018), our reconstruction is conducted at the grid level by an improved calibration method, which could search proxies for a target grid by a new approach of GLDD from its connected areas within a sub-region having homogeneous rainfall regimes and similar precipitation variability. Meanwhile, many new proxies were used, mainly including additional 113 tree-ring width chronologies in the monsoon Asia and more than 6100 dry–wet grade data (23.7 %) from historical documents in China. These additional proxies evidently improve the coverage and distribution of proxies and their temporal homogeneity due to the reconstructed period being limited to 300 years only. This dataset is the first SPI reconstruction covering the whole of Asia based on pure proxies (without long-term observations or climate model constraints) and can be used to more clearly investigate the Asian precipitation change since 1700 and to test the paleoclimate simulation in the industrial period.

https://doi.org/10.5194/essd-14-5717-2022

Earth Syst. Sci. Data, 14, 5717–5735, 2022

Review statement. This paper was edited by Hao Shi and reviewed by four anonymous referees.

References

Adamson, G. C. D. and Nash, D. J.: Documentary reconstruction of monsoon rainfall variability over western India, 1781–1860, Clim. Dynam., 42, 749–769, https://doi.org/10.1007/s00382-013-1825-6, 2014.

Akkemik, Ü., Köse, N., Kopabayeva, A., and Mazarzhanova, K.: October to July precipitation reconstruction for Burabai region (Kazakhstan) since 1744, Int. J. Biometeorol., 64, 803–813, https://doi.org/10.1007/s00484-020-01870-8, 2020.

Altman, J.: Tree-ring-based disturbance reconstruction in interdisciplinary research: Current state and future directions, Dendrochronologia, 63, 125733, https://doi.org/10.1016/j.dendro.2020.125733, 2020.

Altman, J., Fibich, P., Dolezal, J., and Aakala, T.: TRADER: A package for Tree Ring Analysis of Disturbance Events in R, Dendrochronologia, 32, 107–112, https://doi.org/10.1016/j.dendro.2014.01.004, 2014.

Arsalani, M., Azizi, G., and Bräuning, A.: Dendroclimatic reconstruction of May-June maximum temperatures in the central Zagros Mountains, western Iran, Int. J. Climatol., 35, 408–416, https://doi.org/10.1002/joc.3988, 2015.

Arsalani, M., Pourtahmasi, K., Azizi, G., Bräuning, A., and Mohammadi, H.: Tree-ring based December–February precipitation reconstruction in the southern Zagros Mountains, Iran, Dendrochronologia, 49, 45–56, https://doi.org/10.1016/j.dendro.2018.03.002, 2018.

Awan, J. A., Bae, D. H., and Kim, K. J.: Identification and trend analysis of homogeneous rainfall zones over the East Asia monsoon region, Int. J. Climatol., 35, 1422–1433, https://doi.org/10.1002/joc.4066, 2015.

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B., Hoskins, B., and Kurths, J.: Complex networks reveal global pattern of extreme-rainfall teleconnections, Nature, 566, 373–377, https://doi.org/10.1038/s41586-018-0872-x, 2019.

Bombardi, R. J., Kinter, J. L., and Frauenfeld, O. W.: A Global Gridded Dataset of the Characteristics of the Rainy And Dry Seasons, B. Am. Meteorol. Soc., 100, 1315–1328, https://doi.org/10.1175/bams-d-18-0177.1, 2019.

Borgaonkar, H. P., Sikder, A. B., Ram, S., and Pant, G. B.: El Niño and related monsoon drought signals in 523-year-long ring width records of teak (*Tectona grandis* L.F.) trees from south India, Palaeogeography, Palaeoclimatology, Palaeoecology, 285, 74–84, https://doi.org/10.1016/j.palaeo.2009.10.026, 2010.

Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Jones, P. D., Shiyatov, S. G., and Vaganov, E. A.: Tree-ring width and density data around the Northern Hemisphere: Part 1, local and regional climate signals, Holocene, 12, 737–757, 2002.

Buckley, B. M., Stahle, D. K., Luu, H. T., Wang, S. Y. S., Nguyen, T. Q. T., Thomas, P., Le, C. N., Ton, T. M., Bui, T. H., and Nguyen, V. T.: Central Vietnam climate over the past five centuries from cypress tree rings, Clim. Dynam., 48, 3707–3723, https://doi.org/10.1007/s00382-016-3297-y, 2017.

CAMS: Yearly charts of dryness/wetness for the last 500-year period, SinoMaps Press, Beijing, 1981 (in Chinese).

Chen, F., Mambetov, B., Maisupova, B., and Kelgenbayev, N.: Drought variations in Almaty (Kazakhstan) since AD 1785 based on spruce tree rings, Stoch. Environ. Res. Risk Assess., 31, 2097–2105, https://doi.org/10.1007/s00477-016-1290-y, 2016.

Chen, M., Xie, P., Janowiak, J. E., and Arkin, P. A.: Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations, J. Hydrometeorol., 3, 249–266, https://doi.org/10.1175/1525-7541(2002)003<0249:Glpaym>2.0.Co;2, 2002.

Christiansen, B. and Ljungqvist, F. C.: Challenges and perspectives for large-scale temperature reconstructions of the past two millennia, Rev. Geophys., 55, 40–96, https://doi.org/10.1002/2016rg000521, 2017.

Conroy, J. L. and Overpeck, J. T.: Regionalization of Present-Day Precipitation in the Greater Monsoon Region of Asia, J. Climate, 24, 4073–4095, https://doi.org/10.1175/2011jcli4033.1, 2011.

Cook, E. R., Briffa, K. R., and Jones, P. D.: Spatial regression methods in dendroclimatology: A review and comparison of two techniques, Int. J. Climatol., 14, 379–402, https://doi.org/10.1002/joc.3370140404, 1994.

Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., and Wright, W. E.: Asian Monsoon Failure and Megadrought During the Last Millennium, Science, 328, 486–489, https://doi.org/10.1126/science.1185188, 2010a.

Cook, E. R., Seager, R., Heim, R. R., Vose, R. S., Herweijer, C., and Woodhouse, C.: Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context, J. Quat. Sci., 25, 48–61, https://doi.org/10.1002/jqs.1303, 2010b.

Cook, E. R., Seager, R., Kushnir, Y., Briffa, K. R., Büntgen, U., Frank, D., Krusic, P. J., Tegel, W., van der Schrier, G., Andreu-Hayles, L., Baillie, M., Baittinger, C., Bleicher, N., Bonde, N., Brown, D., Carrer, M., Cooper, R., Čufar, K., Dittmar, C., Esper, J., Griggs, C., Gunnarson, B., Günther, B., Gutierrez, E., Haneca, K., Helama, S., Herzig, F., Heussner, K.-U., Hofmann, J., Janda, P., Kontic, R., Köse, N., Kyncl, T., Levanič, T., Linderholm, H., Manning, S., Melvin, T. M., Miles, D., Neuwirth, B., Nicolussi, K., Nola, P., Panayotov, M., Popa, I., Rothe, A., Seftigen, K., Seim, A., Svarva, H., Svoboda, M., Thun, T., Timonen, M., Touchan, R., Trotsiuk, V., Trouet, V., Walder, F., Ważny, T., Wilson, R., and Zang, C.: Old World megadroughts and pluvials during the Common Era, Science Advances, 1, e1500561, https://doi.org/10.1126/sciadv.1500561, 2015.

Cook, E. R., Solomina, O., Matskovsky, V., Cook, B. I., Agafonov, L., Berdnikova, A., Dolgova, E., Karpukhin, A., Knysh, N., Kulakova, M., Kuznetsova, V., Kyncl, T., Kyncl, J., Maximova, O., Panyushkina, I., Seim, A., Tishin, D., Dotny, T. W. O., and Yermokhin, M.: The European Russia Drought Atlas (1400–2016 CE), Clim. Dynam., 54, 2317–2335, 2020.

Coulthard, B. L., St. George, S., and Meko, D. M.: The limits of freely-available tree-ring chronologies, Quat. Sci. Rev., 234, 106264, https://doi.org/10.1016/j.quascirev.2020.106264, 2020.

CRED and UNISDR: The Human Cost of Weather-Related Disasters 1995–2015, Centre for Research on the Epidemiology of Disasters and United Nations International Strategy for Disaster Reduction, Brussels and Geneva, 30, https://www.undrr.org/publication/human-cost-weather-related-disasters-1995-2015 (last access: 22 May 2022), 2015.

Earth Syst. Sci. Data, 14, 5717–5735, 2022

https://doi.org/10.5194/essd-14-5717-2022

Feng, S., Hu, Q., Wu, Q. R., and Mann, M. E.: A Gridded Reconstruction of Warm Season Precipitation for Asia Spanning the Past Half Millennium, J. Climate, 26, 2192–2204, https://doi.org/10.1175/JCLI-D-12-00099.1, 2013.

Friedman, J. H.: A variable span smoother, Technical Report, Department of Statistics, Stanford University, Stanford, CASLAC-PUB-3477, 30, https://doi.org/10.2172/1447470, 1984.

George, S. S.: An overview of tree-ring width records across the Northern Hemisphere, Quat. Sci. Rev., 95, 132–150, https://doi.org/10.1016/j.quascirev.2014.04.029, 2014.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M., and Zhai, P. M.: Observations: Atmosphere and Surface, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 159–254, https://doi.org/10.1017/CBO9781107415324.008, 2013.

Hsu, H. H., Zhou, T. J., and Matsumoto, J.: East Asian, Indochina and Western North Pacific Summer Monsoon – An update, Asia-Pac. J. Atmos. Sci., 50, 45–68, https://doi.org/10.1007/s13143-014-0027-4, 2014.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, J. Climate, 30, 8179–8205, https://doi.org/10.1175/jcli-d-16-0836.1, 2017.

Kamiguchi, K., Arakawa, O., Kitoh, A., Yatagai, A., Hamada, A., and Yasutomi, N.: Development of APHRO_JP, the first Japanese high-resolution daily precipitation product for more than 100 years, Hydrol. Res. Lett., 4, 60–64, https://doi.org/10.3178/hrl.4.60, 2010.

Kostyakova, T. V., Touchan, R., Babushkina, E. A., and Belokopytova, L. V.: Precipitation reconstruction for the Khakassia region, Siberia, from tree rings, Holocene, 28, 377–385, https://doi.org/10.1177/0959683617729450, 2017.

Kucherov, S. E.: Reconstruction of summer precipitation in the Southern Urals over the last 375 years based on analysis of radial increment in the Siberian larch, Russ. J. Ecol., 41, 284–292, https://doi.org/10.1134/s1067413610040028, 2010.

Lee, J., Perera, D., Glickman, T., and Taing, L.: Water-related disasters and their health impacts: A global review, Progress in Disaster Science, 8, 100123, https://doi.org/10.1016/j.pdisas.2020.100123, 2020.

Lever, J., Krzywinski, M., and Altman, N.: Model selection and overfitting, Nat. Methods, 13, 703–704, https://doi.org/10.1038/nmeth.3968, 2016.

Li, Z., Liu, G., Fu, B., Zhang, Q., Hu, C., and Luo, S.: Influence of different detrending methods on climate signal in tree-ring chronologies in Wolong National Natural Reserve, western Sichuan, China, Chinese Journal of Plant Ecology, 35, 707–721, 2011 (in Chinese).

Lian, T., Chen, D., Tang, Y., and Jin, B.: A theoretical investigation of the tropical Indo-Pacific tripole mode, Science China

Earth Sciences, 57, 174–188, https://doi.org/10.1007/s11430-013-4762-7, 2013.

Liu, F., Gao, C., Chai, J., Robock, A., Wang, B., Li, J., Zhang, X., Huang, G., and Dong, W.: Tropical volcanism enhanced the East Asian summer monsoon during the last millennium, Nat. Commun., 13, 3429, https://doi.org/10.1038/s41467-022-31108-7, 2022.

Liu, Y., Hao, Z. X., Zhang, X. Z., and Zheng, J. Y.: Intercomparisons of multiproxy-based gridded precipitation datasets in Monsoon Asia: cross-validation and spatial patterns with different phase combinations of multidecadal oscillations, Climatic Change, 165, 31, https://doi.org/10.1007/s10584-021-03072-6, 2021.

Liu, Y., Zheng, J., Hao, Z., and Ge, Q.: A dataset of standard precipitation index reconstructed from multi-proxies over Asia for the past 300 years, Science Data Bank [data set], https://doi.org/10.57760/sciencedb.01829, 2022.

McCarroll, D., Young, G. H. F., and Loader, N. J.: Measuring the skill of variance-scaled climate reconstructions and a test for the capture of extremes, Holocene, 25, 618–626, https://doi.org/10.1177/0959683614565956, 2015.

Melvin, T. M., Briffa, K. R., Nicolussi, K., and Grabner, M.: Time-varying-response smoothing, Dendrochronologia, 25, 65–69, https://doi.org/10.1016/j.dendro.2007.01.004, 2007.

Murata, A.: Reconstruction of rainfall variation of the Baiu in historical times, in: Climate since AD 1500, edited by: Bradley, R. S., and Jones, P. D., Routledge, London and New York, 224–245, ISBN 0-415-12030-6, 1992.

Nguyen, H. T. T., Turner, S. W. D., Buckley, B. M., and Galelli, S.: Coherent Streamflow Variability in Monsoon Asia Over the Past Eight Centuries – Links to Oceanic Drivers, Water Resour. Res., 56, e2020WR027883, https://doi.org/10.1029/2020wr027883, 2020.

Nieto, R., Ciric, D., Vázquez, M., Liberato, M. L. R., and Gimeno, L.: Contribution of the main moisture sources to precipitation during extreme peak precipitation months, Adv. Water Resour., 131, 103385, https://doi.org/10.1016/j.advwatres.2019.103385, 2019.

North, G. R., Bell, T. L., Cahalan, R. F., and Moeng, F. J.: Sampling Errors in the Estimation of Empirical Orthogonal Functions, Mon. Weather Rev., 110, 699–706, https://doi.org/10.1175/1520-0493(1982)110<0699:Seiteo>2.0.Co;2, 1982.

Palmer, J. G., Cook, E. R., Turney, C. S. M., Allen, K., Fenwick, P., Cook, B. I., O'Donnell, A., Lough, J., Grierson, P., and Baker, P.: Drought variability in the eastern Australia and New Zealand summer drought atlas (ANZDA, CE 1500–2012) modulated by the Interdecadal Pacific Oscillation, Environ. Res. Lett., 10, 124002, https://doi.org/10.1088/1748-9326/10/12/124002, 2015.

Peng, D. D., Zhou, T. J., and Zhang, L. X.: Moisture Sources Associated with Precipitation during Dry and Wet Seasons over Central Asia, J. Climate, 33, 10755–10771, https://doi.org/10.1175/Jcli-D-20-0029.1, 2020.

Pumijumnong, N., Brauning, A., Sano, M., Nakatsuka, T., Muangsong, C., and Buajan, S.: A 338-year tree-ring oxygen isotope record from Thai teak captures the variations in the Asian summer monsoon system, Sci. Rep., 10, 8966, https://doi.org/10.1038/s41598-020-66001-0, 2020.

Qin, D., Hou, S., Zhang, D., Ren, J., Kang, S., Mayewski, P. A., and Wake, C. P.: Preliminary results from the chemical records

of an 80.4 m ice core recovered from East Rongbuk Glacier, Qomolangma (Mount Everest), Himalaya, Ann. Glaciol., 35, 278–284, https://doi.org/10.3189/172756402781816799, 2002.

Sass-Klaassen, U., Leuschner, H. H., Buerkert, A., and Helle, G.: Tree-ring analysis of Juniperus excelsa from the northern Oman mountains, in: Proceedings of the Dendrosymposium 2007, Riga, Latvia, 3–6 May 2007, 99–108, https://doi.org/10.2312/GFZ.b103-08056, 2007.

Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., and Becker, A.: Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC), Atmosphere, 8, 52, https://doi.org/10.3390/atmos8030052, 2017.

Sekaran, U.: Research methods for business: a skill-building approach, 4th edn., John Wiley & Sons, New York, ISBN 0-471-20366-1, 2003.

Seth, A., Giannini, A., Rojas, M., Rauscher, S. A., Bordoni, S., Singh, D., and Camargo, S. J.: Monsoon Responses to Climate Changes – Connecting Past, Present and Future, Current Climate Change Reports, 5, 63–79, https://doi.org/10.1007/s40641-019-00125-y, 2019.

Shah, S. K., Bhattacharyya, A., and Chaudhary, V.: Reconstruction of June–September precipitation based on tree-ring data of teak (*Tectona grandis* L.) from Hoshangabad, Madhya Pradesh, India, Dendrochronologia, 25, 57–64, https://doi.org/10.1016/j.dendro.2007.02.001, 2007.

Shi, F., Zhao, S., Guo, Z., Goosse, H., and Yin, Q.: Multiproxy reconstructions of May–September precipitation field in China over the past 500 years, Clim. Past, 13, 1919–1938, https://doi.org/10.5194/cp-13-1919-2017, 2017.

Shi, H., Wang, B., Cook, E. R., Liu, J., and Liu, F.: Asian Summer Precipitation over the Past 544 Years Reconstructed by Merging Tree Rings and Historical Documentary Records, J. Climate, 31, 7845–7861, https://doi.org/10.1175/Jcli-D-18-0003.1, 2018.

Sinha, A., Berkelhammer, M., Stott, L., Mudelsee, M., Cheng, H., and Biswas, J.: The leading mode of Indian Summer Monsoon precipitation variability during the last millennium, Geophys. Res. Lett., 38, L15703, https://doi.org/10.1029/2011gl047713, 2011.

Stahle, D. W., Cook, E. R., Burnette, D. J., Torbenson, M. C. A., Howard, I. M., Griffin, D., Diaz, J. V., Cook, B. I., Williams, A. P., Watson, E., Sauchyn, D. J., Pederson, N., Woodhouse, C. A., Pederson, G. T., Meko, D., Coulthard, B., and Crawford, C. J.: Dynamics, Variability, and Change in Seasonal Precipitation Reconstructions for North America, J. Climate, 33, 3173–3195, https://doi.org/10.1175/jcli-d-19-0270.1, 2020.

Steiger, N. J., Smerdon, J. E., Cook, E. R., and Cook, B. I.: A reconstruction of global hydroclimate and dynamical variables over the Common Era, Sci. Data, 5, 180086, https://doi.org/10.1038/sdata.2018.86, 2018.

Sun, Q. H., Miao, C. Y., Duan, Q. Y., Ashouri, H., Sorooshian, S., and Hsu, K. L.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, Rev. Geophys., 56, 79–107, https://doi.org/10.1002/2017RG000574, 2018.

Tao, S., Fu, C., Zeng, Z., and Zhang, Q.: Two Long-Term Instrumental Climatic Data Bases of the People's Republic of China (NDP039), ESS-DIVE [data set], https://doi.org/10.3334/CDIAC/cli.ndp039, 1997.

Thompson, L. G., Yao, T., Mosley-Thompson, E., Davis, M. E., Henderson, K. A., and Lin, P. N.: A High-Resolution Millennial Record of the South Asian Monsoon from Himalayan Ice Cores, Science, 289, 1916–1919, https://doi.org/10.1126/science.289.5486.1916, 2000.

Ukhvatkina, O., Omelko, A., Kislov, D., Zhmerenetsky, A., Epifanova, T., and Altman, J.: Tree-ring-based spring precipitation reconstruction in the Sikhote-Alin' Mountain range, Clim. Past, 17, 951–967, https://doi.org/10.5194/cp-17-951-2021, 2021.

Vaganov, E. A., Anchukaitis, K. J., and Evans, M. N.: How Well Understood Are the Processes that Create Dendroclimatic Records? A Mechanistic Model of the Climatic Control on Conifer Tree-Ring Growth Dynamics, vol. 11, in: Dendroclimatology, edited by: Hughes, M. K., Swetnam, T. W., and Diaz, H. F., Springer, Dordrecht, Netherlands, 37–75, https://doi.org/10.1007/978-1-4020-5725-0_3, 2011.

Wang, B., Biasutti, M., Byrne, M. P., Castro, C., Chang, C. P., Cook, K., Fu, R., Grimm, A. M., Ha, K. J., Hendon, H., Kitoh, A., Krishnan, R., Lee, J. Y., Li, J. P., Liu, J., Moise, A., Pascale, S., Roxy, M. K., Seth, A., Sui, C. H., Turner, A., Yang, S., Yun, K. S., Zhang, L. X., and Zhou, T. J.: Monsoons Climate Change Assessment, B. Am. Meteorol. Soc., 102, E1–E19, https://doi.org/10.1175/BAMS-D-19-0335.1, 2021.

Wang, S. and Zhao, Z.: An analyses of historical data of droughts and floods in last 500 years in China, Acta Geographica Sinica, 34, 329–341, https://doi.org/10.11821/xb197904005, 1979 (in Chinese).

Wei, K., Ouyang, C., Duan, H., Li, Y., Chen, M., Ma, J., An, H., and Zhou, S.: Reflections on the Catastrophic 2020 Yangtze River Basin Flooding in Southern China, The Innovation, 1, 100038, https://doi.org/10.1016/j.xinn.2020.100038, 2020.

Wettstein, J. J., Littell, J. S., Wallace, J. M., and Gedalof, Z. e.: Coherent Region-, Species-, and Frequency-Dependent Local Climate Signals in Northern Hemisphere Tree-Ring Widths, J. Climate, 24, 5998–6012, https://doi.org/10.1175/2011JCLI3822.1, 2011.

Wu, R.: Relationship between Indian and East Asian summer rainfall variations, Adv. Atmos. Sci., 34, 4–15, https://doi.org/10.1007/s00376-016-6216-6, 2016.

Xu, C., Sano, M., and Nakatsuka, T.: A 400-year record of hydroclimate variability and local ENSO history in northern Southeast Asia inferred from tree-ring δ^{18}O, Palaeogeogr. Palaeocl., 386, 588–598, https://doi.org/10.1016/j.palaeo.2013.06.025, 2013.

Xu, C., Pumijumnong, N., Nakatsuka, T., Sano, M., and Li, Z.: A tree-ring cellulose δ^{18}O-based July–October precipitation reconstruction since AD 1828, northwest Thailand, J. Hydrol., 529, 433–441, https://doi.org/10.1016/j.jhydrol.2015.02.037, 2015.

Zhang, D. and Liu, C.: Continuation (1980–1992) of the "Yearly Charts of Dryness/ Wetness in China for the last 500-Years Period", Meteorological Monthly, 19, 41–45, 1993 (in Chinese).

Zhang, D., Li, X., and Liang, Y.: Continuation (1992–2000) of the "Yearly Charts of Dryness/Wetness in China for the Last 500-Years Period", Quarterly Journal of Applied Meteorology, 14, 379–388, 2003 (in Chinese).

Zhang, P. Y.: Climate Change in China during Historical Times, Shandong Science and Technology Press, Jinan, ISBN 7-5331-1823-5, 1996 (in Chinese).

Zhang, R. B., Shang, H. M., Yu, S. L., He, Q., Yuan, Y. J., Bolatov, K., and Mambetov, B. T.: Tree-ring-based precipitation re-

Earth Syst. Sci. Data, 14, 5717–5735, 2022

https://doi.org/10.5194/essd-14-5717-2022

construction in southern Kazakhstan, reveals drought variability since AD 1770, Int. J. Climatol., 37, 741–750, 2017.

Zhang, Y. and Wang, K.: Global precipitation system size, Environ. Res. Lett., 16, 054005, https://doi.org/10.1088/1748-9326/abf394, 2021.

Zheng, J., Wu, M., Hao, Z., and Zhang, X.: Spatial pattern of decadal variation of summer precipitation in Eastern China: Comparison of observation and CESM control simulation, Geogr. Res., 35, 14–24, https://doi.org/10.11821/dlyj201601002, 2016 (in Chinese).