

## Article

# Multi-Proxy Reconstruction of Drought Variability in China during the Past Two Millennia

Bing Yang<sup>1,2,3</sup>, Chengguang Lai<sup>4</sup>, Xiaohong Chen<sup>2,\*</sup>, Vijay P. Singh<sup>3</sup> and Jiawen Wang<sup>2</sup>

<sup>1</sup> Guangdong Provincial Science and Technology Collaborative Innovation Center for Water Safety, Guangdong Research Institute of Water Resources and Hydropower, Guangzhou 510635, China; yangb56@mail2.sysu.edu.cn

<sup>2</sup> Center for Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China; wangjw38@mail2.sysu.edu.cn

<sup>3</sup> Department of Biological and Agricultural Engineering, Texas A&M University, College Station, TX 77843, USA; vsingh@tamu.edu

<sup>4</sup> School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510641, China; laichg@scut.edu.cn

\* Correspondence: eescxh@mail.sysu.edu.cn

**Abstract:** Drought imposes serious challenges to ecosystems and societies and has plagued mankind throughout the ages. To understand the long-term trend of drought in China, a series of annual self-calibrating Palmer drought severity indexes (scPDSI), which is a semi-physical drought index based on the land surface water balance, were reconstructed during AD 56~2000. Multi-proxy records of tree-ring width and stalagmite oxygen isotope  $\delta^{18}\text{O}$  were used for this reconstruction, along with random forest regression. The spatiotemporal characteristics of the reconstruction results were analyzed, and comparisons were made with previous studies. Results showed that (1) China witnessed a drought-based state during the past 2000 years (mean value of scPDSI was  $-0.3151$ ), with an average annual drought area of  $85,000 \text{ km}^2$ ; 4 wetting periods, i.e., the Han Dynasty (AD 56~220), the Tang Dynasty (AD 618~907), the Ming Dynasty (AD 1368~1644), and the Qing Dynasty (AD 1644~1912); and 2 drying periods, i.e., the Era of Disunity (AD 221~580) and the Song Dynasty (AD 960~1279). (2) Three different alternating fluctuation dry-wet modes (i.e., interannual, multidecadal, and centennial scales) in China were all significantly ( $p$ -value  $< 0.001$ ) correlated with the amplitude and frequency of temperature in the Northern Hemisphere. (3) According to the spatial models disassembled from the rotated empirical orthogonal function, China was divided into nine dry-wet regions: northwestern China, Xinjiang, southwestern China, southeastern China, the Loess plateau, central China, southwestern Tibet, eastern China, and northeastern China. (4) The random forest (RF) was found to be accurate and stable for the reconstruction of drought variability in China compared with linear regression.

**Keywords:** self-calibrating Palmer drought severity index; meteorological drought; reconstruction; proxy record; multi-scale variation; rotated empirical orthogonal function



**Citation:** Yang, B.; Lai, C.; Chen, X.; Singh, V.P.; Wang, J. Multi-Proxy Reconstruction of Drought Variability in China during the Past Two Millennia. *Water* **2022**, *14*, 858. <https://doi.org/10.3390/w14060858>

Academic Editor: Leonardo V. Noto

Received: 18 January 2022

Accepted: 6 March 2022

Published: 10 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Drought is a recurrent extreme climate phenomenon that has plagued human civilization throughout history [1–3]. It can last for weeks, months, years, or even centuries, and the spatial extent is usually larger than that of other natural hazards (e.g., floods and hurricanes) [4,5], usually resulting in devastating impacts on agriculture, water supply, energy production, ecosystems, and human lives [6]. Understanding the temporal and spatial characteristics of droughts can help choose the appropriate drought mitigation strategy and evaluate future drought risk [7,8].

Drought events can be characterized and monitored with a wide range of measures [7,9]. Among the measures, drought indices, including the standardized precipitation

index [10], standardized precipitation and evapotranspiration index [11], the normalized difference vegetation index [12], the Palmer drought severity index (PDSI) [13,14], and the self-calibrating Palmer drought severity index (scPDSI) [15], are commonly used methods. Among them, the PDSI is a measure of soil moisture availability that has been extensively used to study drought, particularly as the primary indicator of the severity and extent of a recent drought [16,17]. In addition to precipitation, the PDSI also considers the significant effect of temperature on evapotranspiration [18]. Therefore, it provides a more comprehensive way to assess the impact of climate change on drought. The scPDSI, proposed by Wells et al. [15], is considered as an improved version of the “traditional” PDSI. It calibrates the behavior of the index automatically at any location by replacing the empirical constants derived from climatic characteristics and duration factors with dynamically calculated values [19]. This index has become a popular drought metric in studies that quantify possible trends in future soil moisture availability [17,20,21].

Long-term instrumental records of climate elements generally provide reliable information about drought variability from interannual to decadal scales and beyond. However, most current studies have focused on individual events and/or instrumental records of  $\leq 100$  years [22], due to the limitation of the length of existing instrumental records. Short-term instrumental records/or individual events do not contain the stable characteristics and regularity of variability of drought conditions over a long time scale, and are unfavorable to study the characteristics and driving mechanisms of space-time variability [23]. Whether the duration of the current drought pattern is decadal or a century remains unclear; however, if the latter holds true, local ecological and economic systems must adapt to the changing environmental conditions [24]. A better understanding of such changes in historical time would provide insights into a possible physical mechanism and help predict the direction of the environmental evolution under an expected global warming scenario. Therefore, long-term drought variability derived from high resolution and continuous paleoclimatic proxy data, such as tree rings, stalagmites, ice cores, and historical documents, is invaluable [25]. Among these, annually resolved and precisely-dated tree ring and stalagmite data are most commonly used [26–29]. These two types of data have been used to investigate the paleoclimate throughout the world, including the North American Drought Atlas [30], the Monsoon Asia Drought Atlas [31], the Old-World Drought Atlas [32], and the Australian and New Zealand Summer Drought Atlas [33], May–September precipitation in China over the past 500 years [34], and the six hundred-year annual minimum temperature history for the central Tibetan Plateau [35].

Most previous studies that analyzed long-term drought variability were based on a statistical modeling approach that first established an empirical relationship between climatic proxies (e.g., tree-rings and stalagmites) and drought for the instrumental period [36]; then, they carried out a drought reconstruction by feeding the climatic proxies of the paleo period into the established relationship [37,38]. Among the available approaches, the principal component linear regression (henceforth simply referred to as linear regression) has often been adopted [39,40] and has been shown to provide reliable reconstruction. The obvious disadvantage of linear regression is the poor estimation of the amplitude of the past variations at the lower frequencies. This defect might result in an underestimate (or overestimate) of an actual drought. In this case, a novel alternative method for overcoming these deficiencies is much needed.

The random forest (RF), a flexible machine-learning algorithm proposed by Breiman [41], is a combination regression method, based on statistical learning theory. In the RF, multiple samples are drawn using the resampling bootstrap method, and regression trees are built corresponding to each bootstrap sample. Eventually, all regression trees are combined, and final regression results are obtained by voting. The RF algorithm is a natural, non-linear modelling tool that provides estimates regarding the hierarchy of variables in a regression; and thus can be used to estimate each sample’s contribution to the final results [42]. A great deal of theoretical and empirical studies have turned out to perform very well compared to linear regression, including deeper mining of potential information, higher regression

accuracy, acceptable tolerance to outliers and noise, and robustness against over-fitting [43]. Based on this, RF should theoretically be a better option to obtain a regional pattern of drought in conjunction with multi-proxy records with good spatial coverage, allowing one to fully analyze the dynamic space-time behavior of droughts. However, few applications in this field have been previously reported.

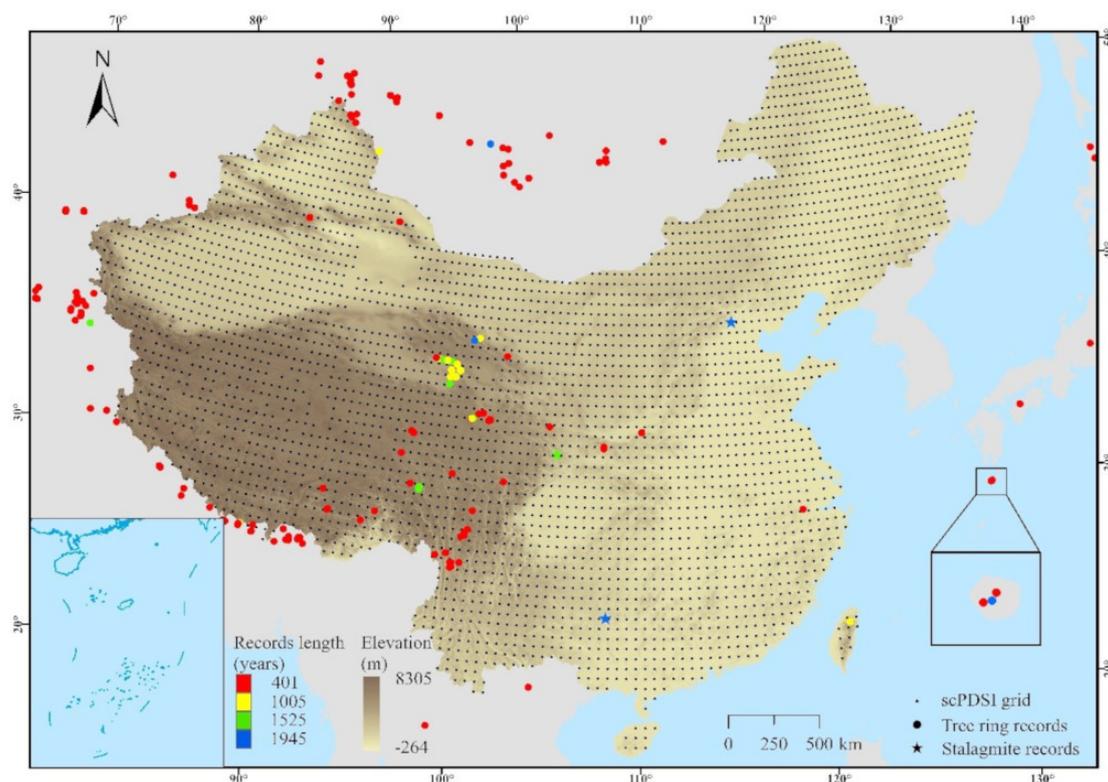
China, with its complicated climatic conditions, large population, fragile ecosystem, and rapidly developing economy, suffers from frequent severe droughts [44]. Major droughts, particularly those that have occurred in the past 20 years have caused large economic losses in China [45]. According to the statistics of the Ministry of Water Resources of China (<http://www.mwr.gov.cn/>, accessed on 17 January 2022), the annual average disaster area impacted by drought during 2011–2016 accounted for 70% of the total area, with an affected population of 16.43 million and economic losses of >12 billion USD. Hence, in the last decade, a great deal of information about drought variation has been collected [21,46–48]. However, studies in China have mainly focused on short-term drought variability with records of <60 years [7,49], due to limited observational data. The only long-sequence drought studies are confined to some individual regions, e.g., arid to semi-arid areas of China [50,51]; the southeastern Tibetan Plateau [52]; southwest China [53]; central Inner Mongolia, China [28]; northwestern China [54]; the northern fringe of the Asian summer monsoon region [55]; and the western Qilian Mountains of northwestern China [54]; the trans-Himalayan region of central Himalaya [29]; and the central Tibetan Plateau [35], while a the long period of historical drought reconstruction in all of China is rare. These studies of fragmented regions across China are not conducive to understanding the drought variability over the long-term on a national scale.

Therefore, the primary objective of this study was to investigate drought variability in China during the past two millennia with multi-proxy reconstruction using RF regression. The investigation included: (1) the collection of multi-proxy records of China for AD 56~2000 and the scPDSI dataset for AD 1951~2000; (2) analysis of the correlation between the proxy records and the scPDSI dataset during the overlap period (AD 1951~2000) to demonstrate the feasibility of reconstruction and build a reconstruction model based on RF regression, and then calibration and verification of the model to ensure the reliability of the results; (3) reconstruction of the scPDSI during AD 56~2000 and analysis of the spatiotemporal characteristics based on the reconstruction results. Finally, we compared previous studies with the linear regression results to prove the rationality of the results. This reconstruction is an important source of information for documenting climate change over the interannual to centennial scale, and the proposed method is expected to be a novel way to analyze drought intensity and the driving mechanism of drought.

## 2. Study Area and Data

### 2.1. Study Area

China lies between latitudes 18° and 54° N and longitudes 73° and 135° E. Located in eastern Asia and on the western shore of the Pacific Ocean, China (Figure 1) has many different types of topography. The terrain is higher in the west and lower in the east, with roughly a three-ladder-like distribution [56]. A temperate monsoon climate predominates in the northeast region, and a temperate continental climate dominates the northwest region. A sub-tropical monsoon climate prevails in Southern China, except for the southernmost corner, where a tropical monsoon climate is dominant. The Qinghai-Tibet Plateau is dominated by a plateau alpine climate. Overall, China is located within the East Asian monsoon region and experiences significant monthly, annual, and inter-annual variability in precipitation and temperature. A semi-arid or semi-humid climate dominates the northern parts of eastern China, with annual precipitation of 200~800 mm, and the southern part has a relatively wetter climate with annual precipitation of 800~2000 mm.



**Figure 1.** Geographical position and territory of China and the proxy records distribution map.

## 2.2. Data

### 2.2.1. The scPDSI Grid

In this study, the scPDSI was calculated for the period AD 1951~2000 based on the Climatic Research Unit (CRU) TS 3.10.01 dataset (3725 grid points,  $0.5^\circ \times 0.5^\circ$  resolution). The dataset was created using the highly quality-controlled station records, which were checked for inconsistencies and adjusted where necessary, but edits caused by a change in observation practices were avoided. Gridded precipitation, temperature, cloud cover, wind speed, and vapor-pressure data were selected from the annually dataset compiled by the CRU (<http://www.badc.rl.ac.uk/>, accessed on 17 January 2022). This dataset was selected because it offers three advantages [17]. First, the scPDSI has a similar range of variability in diverse climates, which makes it an appropriate metric for comparing the relative availability of moisture in different regions. Second, the more physical Penman-Monteith parameterization was used to calculate potential evapotranspiration. Finally, the seasonal snowpack dynamics considered in the water balance model provided a more accurate measure of the availability of moisture for snowy regions when the snowpack melted. A more complete review of the scPDSI dataset is presented by Schrier et al. [17]. The scPDSI values and drought categories are shown in Table 1.

**Table 1.** Classification of the scPDSI Values.

scPDSI Value	scPDSI Category	scPDSI Value	scPDSI Category
Above 4.00	Extreme wet	Below $-4.00$	Extreme drought
3.00 to 3.99	Severe wet	$-3.00$ to $-3.99$	Severe drought
2.00 to 2.99	Moderate wet	$-2.00$ to $-2.99$	Moderate drought
1.00 to 1.99	Mid wet	$-1.00$ to $-1.99$	Mid drought
0.50 to 0.99	Incipient wet	$-0.50$ to $-0.99$	Incipient drought
0.49 to $-0.49$	Normal		

### 2.2.2. Selected Proxy Data and Pre-Processing

In total, we assembled 234 proxy records (data resolution is 1 year) that included 231 tree-ring width chronologies and 3-stalagmite oxygen isotope  $\delta^{18}\text{O}$  datasets. With the purpose of identifying long-timescale climate (drought) signals, 231 tree ring chronologies were obtained from the NOAA Paleoclimatology Program's International Tree Ring Data Bank (ITRDB) [57], and three stalagmite oxygen isotope  $\delta^{18}\text{O}$  datasets were acquired from the website [57]. The tree ring chronology data were contributed by ITRDB, and the stalagmite oxygen isotope  $\delta^{18}\text{O}$  data were contributed by Wang et al. [58] and Tan et al. [59]. The reliability of the data was guaranteed as follows: (1) standard techniques of crossdating and detrending were used to process these proxy records by the data contributor; (2) these proxy data were used in related studies [31,34]. The length of these proxy data ranges between 400 and 1944 years. We synthesized 234 proxy records from China and surrounding areas, as shown in Figure 1. Each proxy record was guaranteed to be significantly correlated with one or more scPDSI records at the 90% confidence level ( $|r| > 0.2329$ ,  $n = 50$ ,  $p$ -value  $< 0.1$ ) during the overlap period. Basic information of the five longest proxy records is shown in Table 2.

**Table 2.** Basic information table of the five longest proxy records.

NO	Country	Area	Site	Lat (° N)	Long (° E)	Archive Type	Proxy Measurement	Reference
1	Japan (JP)	Asia2k	YKS	30.33	130.5	Tree ring	Total ring width	[57]
2	Mongolia (MG)	LDEO	SODAPS	48.3	98.93			
3	China (CH)	CHIN070	HYGJU	38.57	99.33	Speleothem	$\delta^{18}\text{O}$	[58]
4	China (CH)	Dongge Cave	DGC	25.28	108.08			
5	China (CH)	Shihua Cave	SHC	39.78	115.93			

The time span of these five records is all AD 57–2000.

## 3. Methodology

### 3.1. Random Forest

Classification and regression Trees (CART) is a decision tree algorithm for both classification and regression. It was first described by Breiman [41]. It is a recursive algorithm, which partitions the training dataset by doing binary splits. It is a conceptual simple decision tree algorithm, and performs acceptably in practice. CART is probably the closest to having the desired combination of features. It handles high-dimensional data well; has the ability to ignore irrelevant descriptors; handles multiple mechanisms of action; and is amenable to model interpretation. The major drawback, however, is that CART usually has relatively low prediction accuracy. This drawback may impede its use in applications such as the virtual screening of compound libraries. Because of the great appeal of CART, there have been many efforts to improve its prediction accuracy. These attempts resulted in a large number of various tree-based algorithms. It has recently been discovered that one of the best ways to improve the performance of decision tree-based algorithms is to use ensembles of trees. In this paper, we present one such ensemble method, random forest.

RF, proposed by Breiman [41], is an improved classification and regression tree method that has gained popularity for its robustness and flexibility when modeling the input-output functional relationship. Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. Such a method consists of a collection of regression trees trained using different bootstrap samples of the training data [42,60]. The generalization error for forests converges as to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them. Significant improvements in classification accuracy have resulted from growing an ensemble of trees and letting them vote for the most popular class. In order to grow these ensembles, often random vectors are generated that govern the growth of each tree in the ensemble. Each

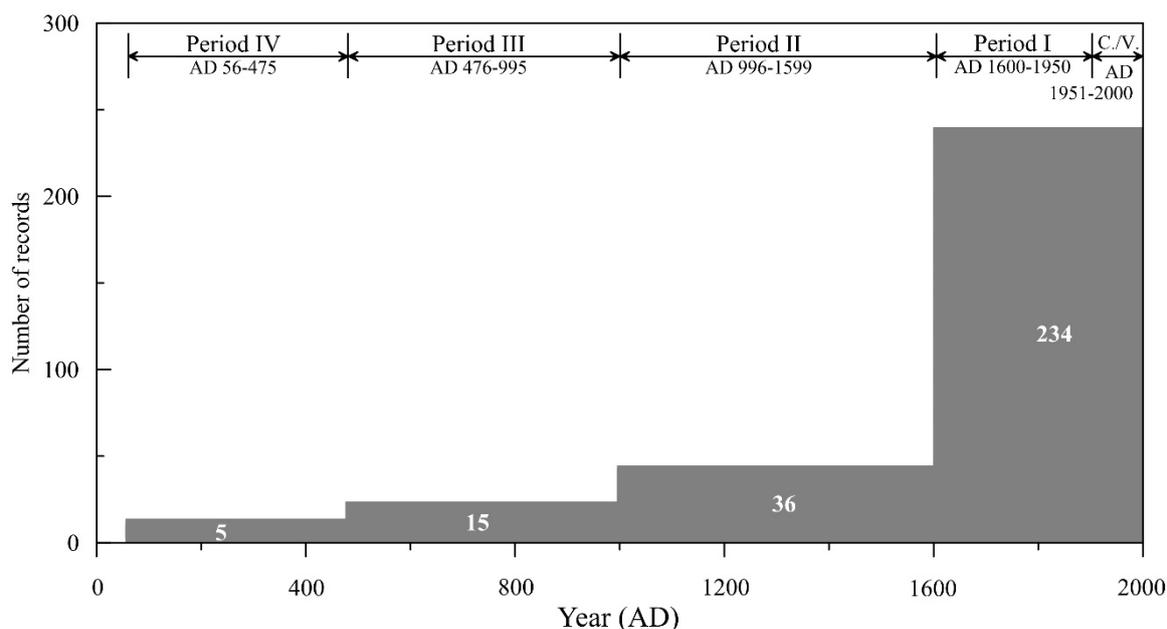
tree acts as a regression function on its own, and the final output is taken as the average of the individual tree outputs. Moreover, due to the RF's built-in cross-validation capability, carried out with the help of out-of-bag samples, it provides realistic prediction error estimates during the training process. After a large number of trees are generated, they vote for the most popular class. We call these procedures random forests.

Random forest regression is a function of randomForest (version 4.6-14), while the randomForest is a package that enables one to carry out classification and regression based on a forest of trees using random inputs in R language (open-source compiler).

### 3.2. Reconstruction Model

Previous studies have repeatedly verified a significant relationship between tree-ring, stalagmite and drought reconstructions [61]. The time stability of the model was tested by using calibration and verification methods [36,62]. To reconstruct the scPDSI, we used a calculated annual scPDSI gridded dataset, which the proxy records were significantly related to scPDSI in the domain studied.

According to the number of candidate proxy records for the different periods, the entire research period was divided into five parts (Figure 2): the calibration/verification period (C./V., AD 1951~2000), reconstruction period I (AD 1600~1950), reconstruction Period II (AD 996~1599), reconstruction period III (AD 476~995), and reconstruction Period IV (AD 56~475). The reconstruction procedure was adopted as follows:



**Figure 2.** Diagram of the candidate proxy records for each year, and five study periods.

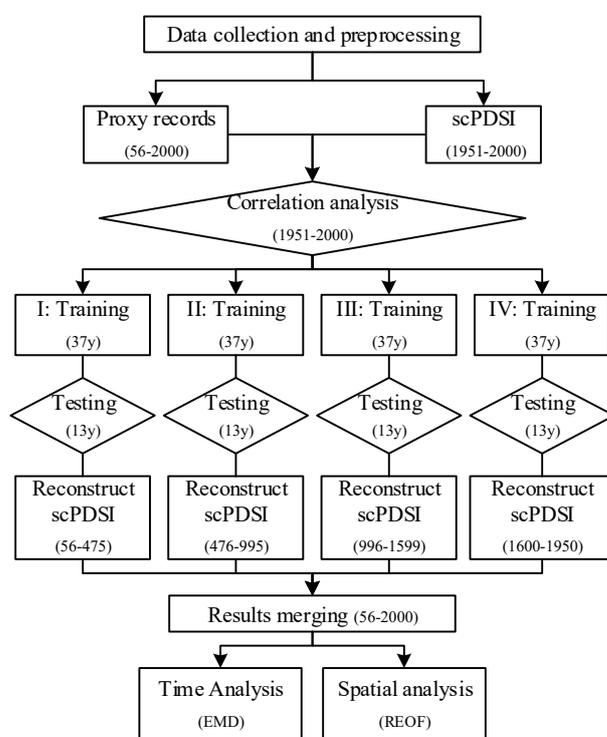
**Step 1: Correlation analysis.** We selected the corresponding proxy records of the 4 reconstruction periods during 1951~2000 to carry out Pearson's correlation analysis with 3725 annual scPDSI, respectively. This was used to illustrate the feasibility of reconstruction.

**Step 2: Modeling and calibration.** The proxy records of the 4 reconstruction periods during 1951~2000 were selected as the independent variable, while the scPDSI was selected as the dependent variable. These independent and dependent variables make up the four model datasets. Then, the scPDSI reconstruction model was constructed by sampling three-quarters of the modeling data (37 years) without random replacement, and input into the RF ( $mtry = 78$ ,  $ntree = 15,000$ ) [42,43]. The prediction results from modeling were correlated with the dependent variables in the three-quarters of the modeling data to calibrate the training accuracy of the model.

**Step 3: Model validation.** The independent variables in the remaining one-quarter of the modeling data (13 years) were inputted into the reconstruction model. The testing

accuracy of the model was verified by correlation analysis of the model prediction results with the dependent variables in one-quarter of the modeling data. The traditional accuracy and skill parameters, including reduction of error (RE) and the coefficient of efficiency (CE) during the verification period [63], were used to evaluate the reliability of reconstructions. The uncertainty of the model was calculated using the standard deviation of the residual between the reconstructed and dependent variables during the verification period. To ensure the accuracy of reconstruction results, we selected the grid points with qualified training accuracy ( $n = 37$ ,  $r > 0.2709$ ,  $p$ -value  $< 0.1$ ) and qualified testing accuracy ( $n = 13$ ,  $r > 0.4575$ ,  $p$ -value  $< 0.1$ ), also called qualified points, to carry out the next scPDSI reconstruction.

Step 4: scPDSI reconstruction. The different proxy records of all qualified points were inputted into the four corresponding scPDSI reconstructions to reconstruct the annual scPDSI. To match the length of all reconstructed data for the 3725 grid points, the Kriging interpolation method was used to calculate the estimated scPDSI for all unqualified points, based on the reconstructed scPDSI values. Finally, the reconstructed and interpolation results were combined to obtain the complete reconstructed scPDSI sequence (AD 56~2000). The detailed flowchart is shown in Figure 3.



**Figure 3.** Flowchart of the reconstruction model.

### 3.3. Empirical Mode Decomposition (EMD)

To fully understand the multi-scale variability of reconstructed scPDSI series, the EMD [64] was used to analyze the reconstructed scPDSI series and four group temperature anomaly sequences. The EMD is based on the direct extraction of the energy associated with various intrinsic time scales. It is a very efficient method to fully analyze and understand a long sequence multi-scale change feature. The advantage is that the signal components of different scales can be preserved and extracted, which helps better understand the multi-scale change characteristics and the general trend in the target sequence. EMD essentially decomposes the fluctuation or trend of different scales (frequencies) from the original signal according to the natural signal oscillation.

A series of intrinsic mode function (IMF) components with different scales was obtained, and all components can be reverted to the original signal. Each IMF corresponded to an scPDSI fluctuation of one band, reflecting the distribution and variability of the main pe-

riod and amplitude in the time domain at different scales. The lowest frequency component (residual IMF) represents the original trend signal term. The trend term represents a linear term in the signal or a slowly changing component with a period greater than the length of the signal data. Each IMF component change law agrees with the non-linear change characteristics of the natural signal, and the fluctuation has a relatively stable quasi-period called the main period. The influence of different scale signals on the overall characteristics of the original data was all different. The square of the amplitude of each component reflected the signal strength and energy of the component in the original data.

### 3.4. Rotated Empirical Orthogonal Function (REOF)

The empirical orthogonal function (EOF) analysis introduced by Lorenz [65] is a convenient and effective method extensively used in atmospheric, oceanic, and climatic research. The empirical orthogonal function analysis identifies and extracts spatiotemporal modes that are ordered by considering their representations of data variance. EOF analysis output includes spatial patterns (EOFs), temporal coefficients (principal components, PCs), and eigenvalues. As an effective eigen method for phenomenon identification and space reduction, EOF analysis is widely used in climate research. However, because of its orthogonality constraint, EOF analysis has a tendency to produce unphysical modes. Previous studies have shown that the drawbacks of EOF analysis could be partly alleviated by rotated EOF (REOF) analysis. REOF analysis was introduced by Richman to simplify physical mechanisms underlying characteristic patterns or to seek “physical” modes [66]. REOF supplies a new set of modes by rotating the vector space of the initial empirical orthogonal function, and improves physical interpretation of the original field. As one of the most popular types of REOF schemes, varimax rotation linearly transforms spatial patterns derived by empirical orthogonal function analysis into a rotated basis (REOF) and relaxes the orthogonality of the spatial patterns based on certain criteria [67].

## 4. Results

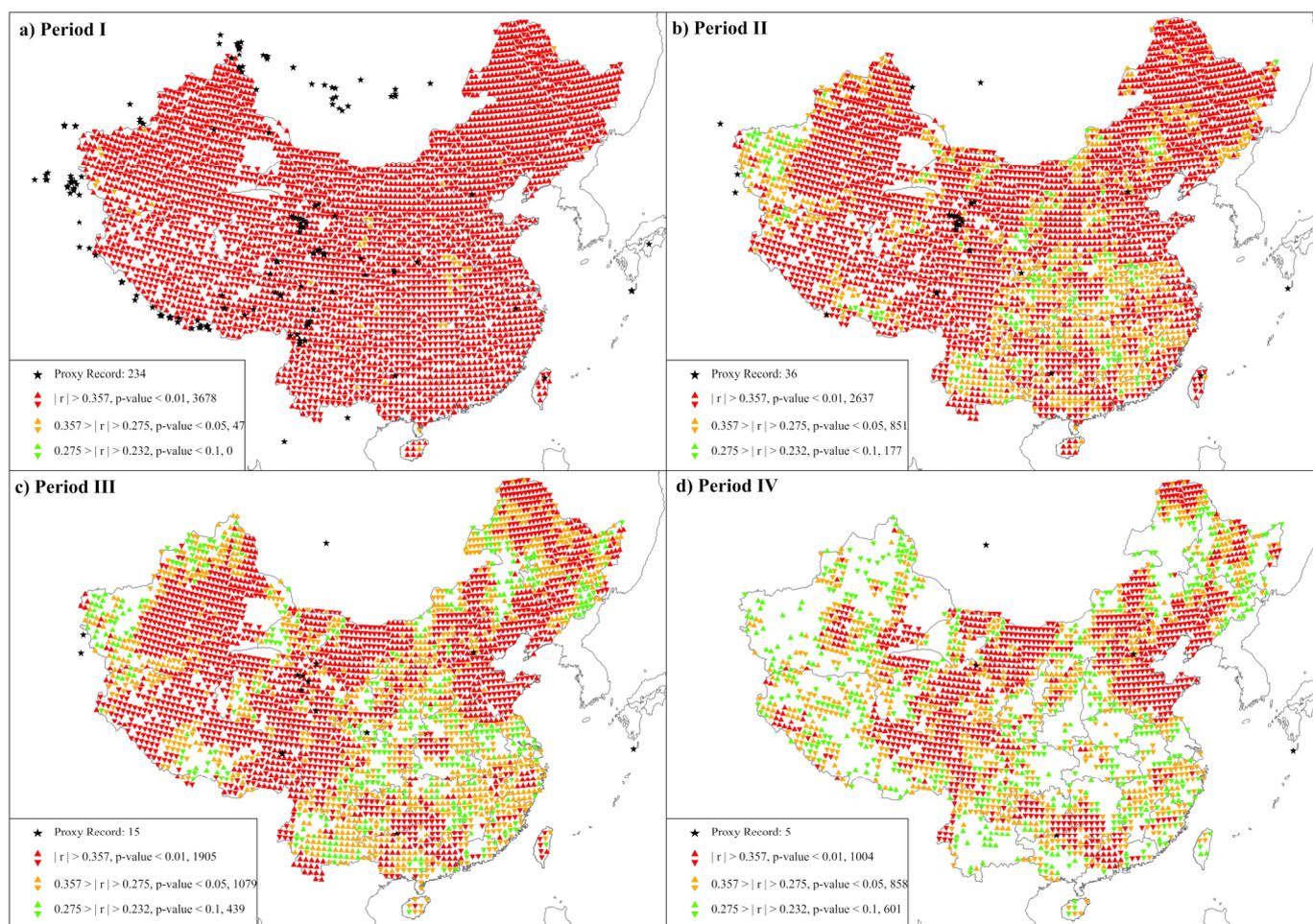
### 4.1. Model Feasibility

#### 4.1.1. Data Correlation

The numbers of candidate proxy records for the 4 reconstruction periods were 234 (Period I), 36 (Period II), 15 (Period III), and 5 (Period IV), respectively. According to step 1 in Section 3.2, the R language was used to perform Pearson’s correlation analysis of the proxy records and scPDSI during AD 1951–2000. To ensure that at least 1 proxy record was significantly correlated ( $|r| > 0.2329$ ,  $n = 50$ ,  $p$ -value  $< 0.1$ ) with one grid point, we selected the maximum correlation coefficient of the grid point and all proxy records as this point’s correlation analysis results (Figure 4). During the 4 periods, 3725 (100%), 3665 (98.39%), 3423 (91.89%), and 2463 (66.12%) grid point correlation coefficients met the reconstruction requirements ( $|r| > 0.2329$ ,  $n = 50$ ,  $p$ -value  $< 0.1$ ), respectively. With a reduction in the number of proxy records, the qualified grid points were gradually reduced. These qualified grid points were evenly distributed even in period IV, indicating that these points had good spatial representation.

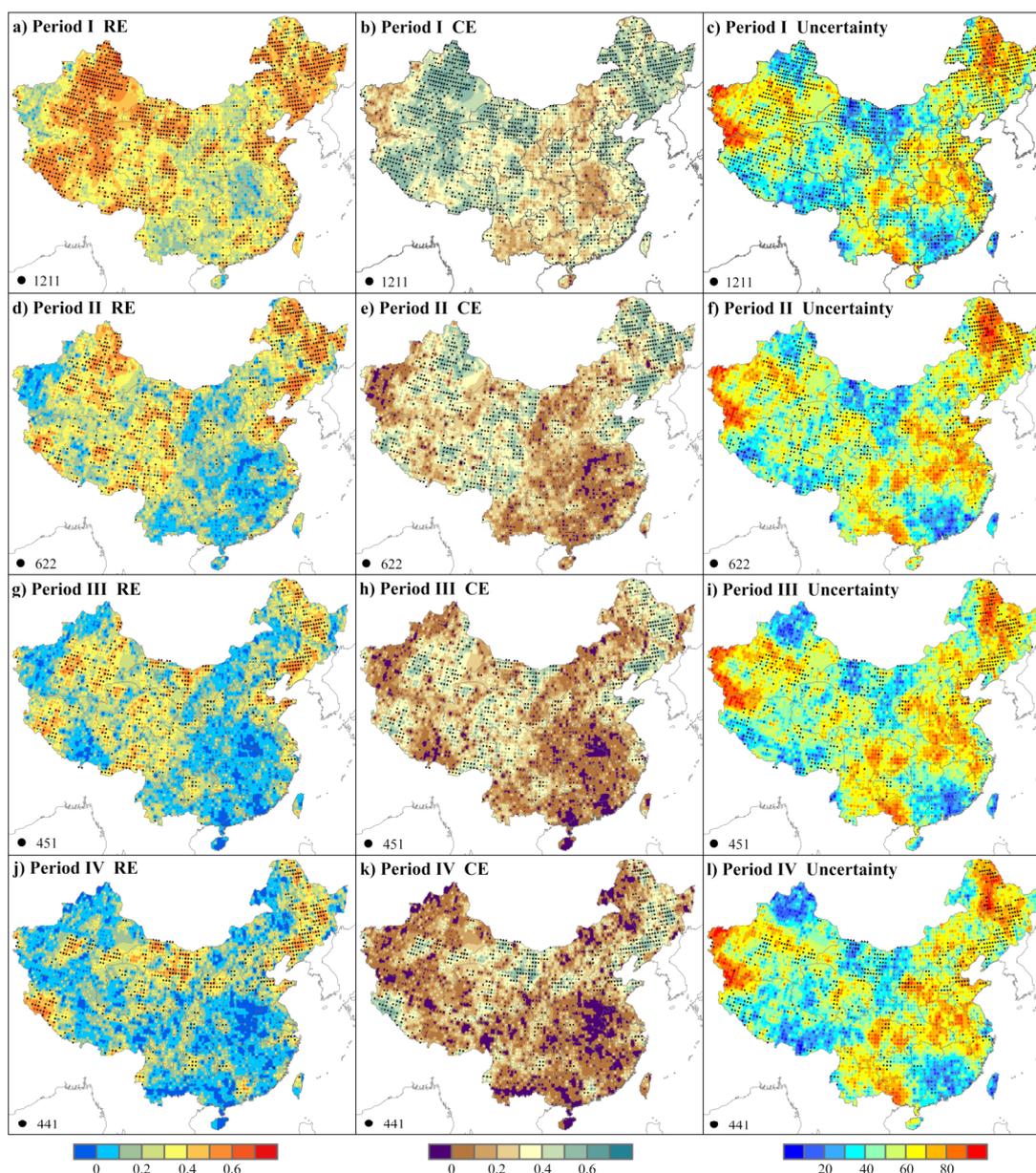
#### 4.1.2. Calibration and Verification

According to steps 2 and 3 in Section 3.2, the reconstruction models of the four periods were constructed, calibrated, and verified. Among them, there were 1211, 622, 451, and, 441 grid points with qualified training accuracy and qualified testing accuracy during the four reconstruction periods, respectively (black dots in Figure 5). These points were evenly distributed and had a particular geographical representation of the regional climate.



**Figure 4.** Results of Pearson correlation analysis of four reconstruction periods between proxy records and scPDSI (0.5° × 0.5° resolution). The normal triangle represents a significant positive correlation, while the inverted triangle represents a significant negative correlation.

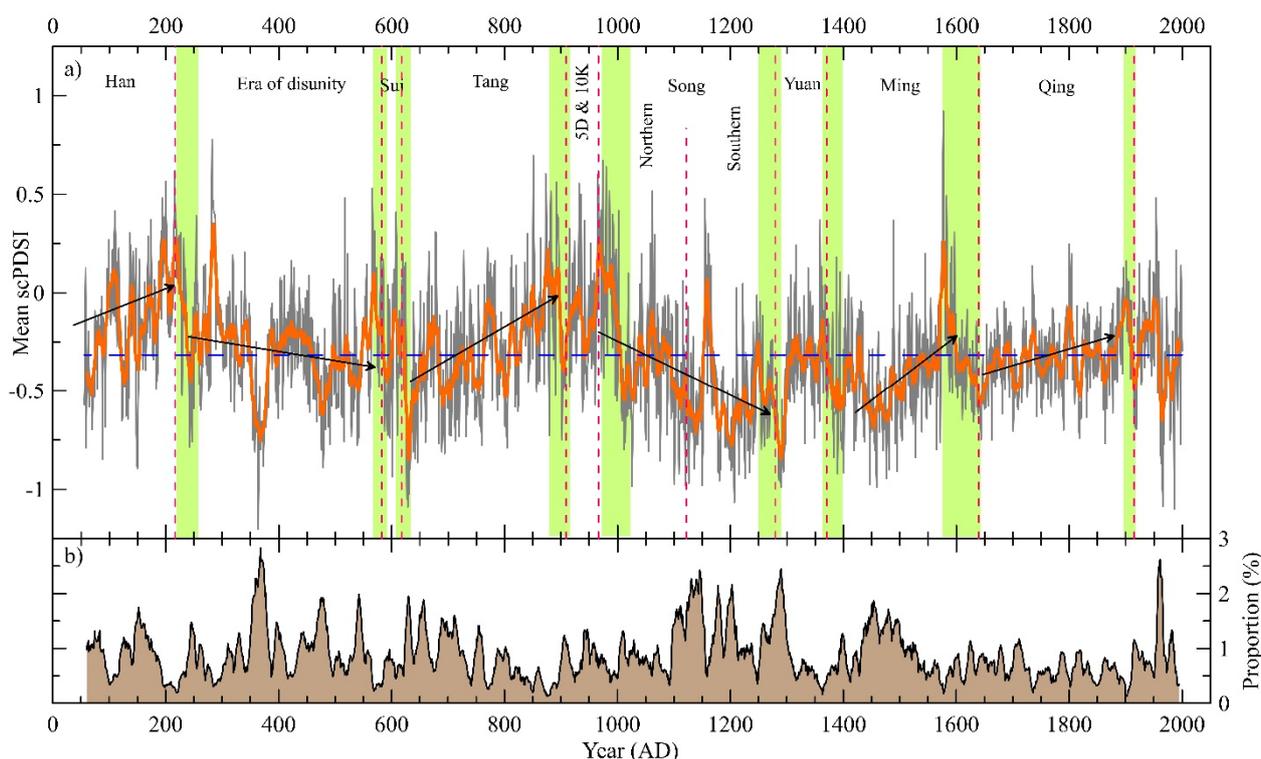
Figure 5 presents a summary of the reconstruction skills. It shows that the similarity in the patterns among the RE and the CE maps during different periods. The western China and northeastern China were characterized by a better quality of reconstruction than the central China and southeastern China. It is natural that the Western China has a better quality of reconstruction because of the wide distribution of proxy records. Although the Northeast China (mainly Changbai Mountain and Daxinganling) lacks proxy records, the results of model validation in this area are statistically reliable. This may be some teleconnection [68], which needs further study, between this region and adjacent proxy records. Positive RE and CE values suggested a robust verification, as CE is generally considered a particularly rigorous indicator of model skill [31]. During the 4 periods, the RE and CE values showed a gradual decreasing trend, but the overall proportion of grid points greater than 0 was above 80%. The trend of uncertainty in the four periods was basically consistent, showing that the model was stable during the different reconstruction periods. At the same time, the high level of uncertainty was mostly concentrated in the region where the proxy records were relatively scarce. In summary, the reconstruction model of this study was reasonable and reliable.



**Figure 5.** Skills of the reconstructed scPDSI in China for the calibration/verification period. The black dots represent grid points with qualified training accuracy ( $n = 37$ ,  $r > 0.2709$ ,  $p$ -value  $< 0.1$ ) and qualified testing accuracy ( $n = 13$ ,  $r > 0.4575$ ,  $p$ -value  $< 0.1$ ). The RE and CE are the reduction of error and the coefficient of efficiency, and the uncertainty is characterized by the standard deviation of the residual between the reconstructed and calculated scPDSI during the verification period.

#### 4.2. Results of the scPDSI Reconstruction

To visualize the regularity of the dry-wet change on an interannual scale, we calculated the average value of annual reconstructed scPDSI (Figure 6a), constituting the mean scPDSI interannual variability curve for China during the past two millennia. Based on the long-term mean value of the reconstructed scPDSI (Figure 6a, the mean value is  $-0.3151$ ), the present state of China is wetter than the past two millennia. We counted the number of grid points that had severe droughts (scPDSI  $\leq -3$ ) every year, and then calculated their proportions as a percentage of the total grid to characterize annual drought severity in all of China (Figure 6b). Among them, the mean value of the proportion of severe drought accounted for 0.89% (35/3725), which converted into an area of about 85,000 km<sup>2</sup> (resolution of the grid was  $0.5^\circ \times 0.5^\circ$ ).



**Figure 6.** (a) The mean scPDSI interannual variation curve of China during the past two millennia (AD 56–2000). The gray line represents the annual mean value of the reconstructed scPDSI of whole of China. The orange line represents the 10-year moving average value of the reconstructed scPDSI of the whole of China. The black dashed line represents the perennial average value of the reconstructed scPDSI of the whole of China. The vertical pink dashed lines are dynasty dividing lines. (b) Proportion diagram of severe drought events ( $\text{scPDSI} \leq -3$ , Table 1).

We define a period for which the scPDSI is below (above) the mean value for more than 10 consecutive years as a significant dry (wet) period. Based on the relationship between the reconstructed scPDSI and its mean value, we extracted 19 significant dry periods and 18 significant wet periods during the past 2000 years (Table 3). As shown in Table 2, there were 13 significant humid periods before AD 1000, and only 5 after AD 1000. There were 9 significant dry periods before AD 1000, and 10 after AD 1000. Combined with the trend of change shown in Figure 6a, the millennium before AD 1000 was more humid than the millennium after AD 1000.

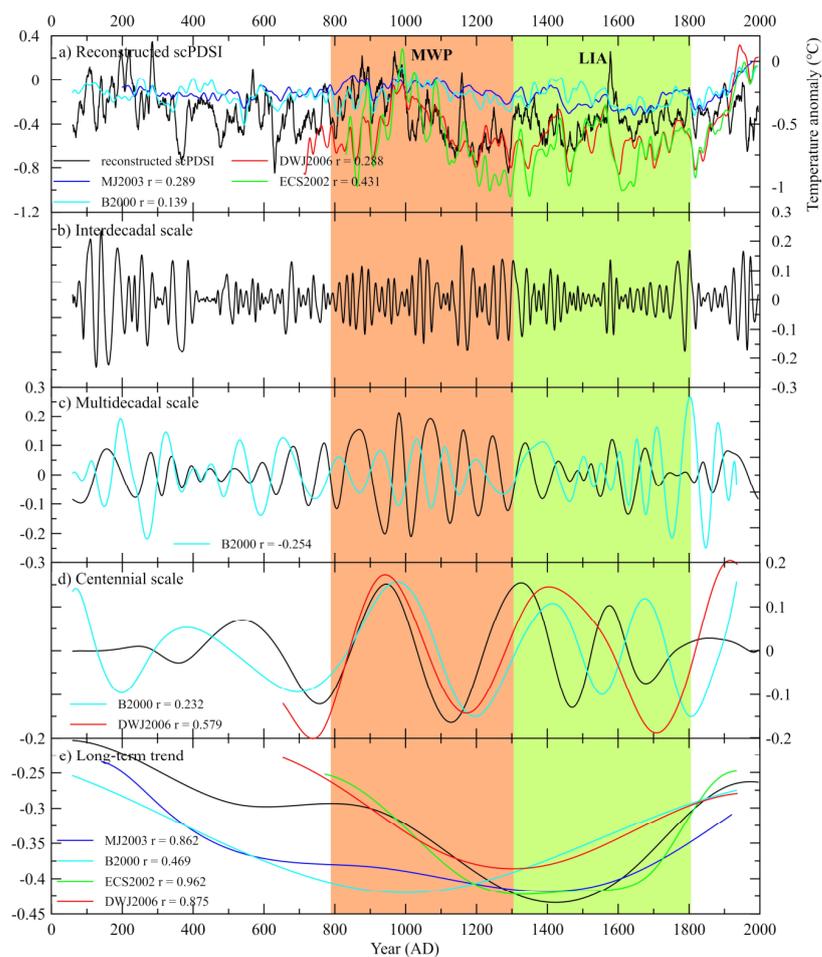
### 4.3. Spatiotemporal Variations of the Reconstructed scPDSI

#### 4.3.1. Multi-Scale Variation

We analyzed the reliability of our reconstructed scPDSI series by correlation analysis with the average annual (MJ2003, Mann and Jones [69]; ECS2002, Esper et al. [70], recalibrated by Cook et al. [71]; DWJ2006, D'Arrigo et al. [72]) or summer (B2000, Briffa, [73]) temperature anomalies in the northern hemisphere (Figure 7a). Significant positive correlations were detected between our reconstructed scPDSI series and these 4 groups of temperature anomalous sequences (Figure 7a,  $r_{\text{MJ2003}} = 0.289$ ,  $n = 1780$ ,  $p\text{-value} < 0.001$ ;  $r_{\text{B2000}} = 0.139$ ,  $n = 1934$ ,  $p\text{-value} < 0.001$ ;  $r_{\text{DWJ2006}} = 0.288$ ,  $n = 1283$ ,  $p\text{-value} < 0.001$ ;  $r_{\text{ECS2002}} = 0.431$ ,  $n = 1162$ ,  $p\text{-value} < 0.001$ ), and the maximum correlation coefficient was 0.431. Analysis showed that the history of dry-wet conditions in China was consistent with the temperature change in the northern hemisphere.

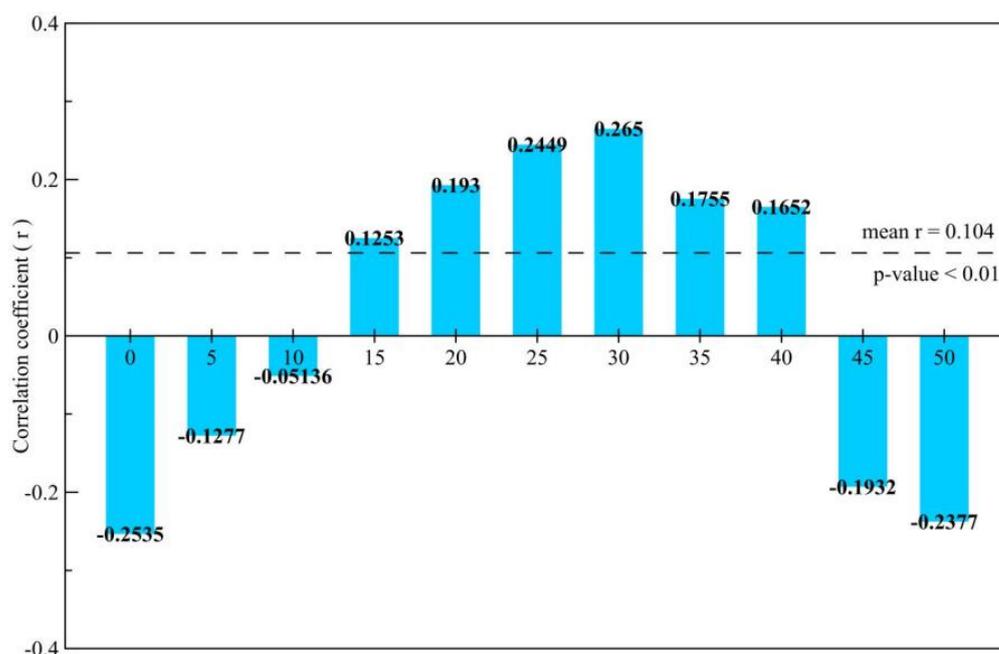
**Table 3.** Statistics of the significant drought/humid periods.

Drought Periods			Humid Periods		
Number	Start & End Year (AD)	Duration (Year)	Number	Start & End Year (AD)	Duration (Year)
1	60–74	15	1	77–87	11
2	122–131	10	2	93–120	28
3	350–381	32	3	133–146	14
4	463–494	32	4	154–238	85
5	524–546	23	5	276–302	27
6	623–650	28	6	335–348	14
7	684–716	33	7	548–584	37
8	744–761	18	8	602–621	20
9	798–807	10	9	765–785	21
10	1007–1031	25	10	809–818	10
11	1096–1151	56	11	828–900	73
12	1169–1247	79	12	910–940	31
13	1254–1302	49	13	958–1002	45
14	1377–1403	27	14	1153–1166	14
15	1433–1508	76	15	1570–1599	30
16	1624–1666	43	16	1794–1807	14
17	1697–1715	19	17	1882–1910	29
18	1812–1823	12	18	1925–1958	34
19	1912–1921	10			



**Figure 7.** Comparison of the reconstructed scPDSI and four groups of temperature anomaly of the northern hemisphere over the past two millennia (AD 56–2000) using the empirical mode decomposition.

Four different IMF components (i.e., interannual scale (<10 years), multidecadal scale (10~100 years), centennial scale (100~1000 years), and long-term trends (>1000 years)) were extracted and are shown in Figure 7. Based on the redfit function (which is part of Schulz's REDFIT (version 3.8e) program and estimates the red-noise spectra of a time series in dpLR, while the dpLR (the Dendrochronology Program Library in R) is a package that enables dendrochronologists to handle data processing and analysis) in the package of "dpLR" in R, red-noise spectral analysis was used to assess the oscillatory domains in these components. The dominant interannual variation was 5~10 years (Figure 7b), which was the strongest signal component in the original sequence. The correlation coefficient between this component and the reconstructed scPDSI series was 0.353, which reflected the main high-frequency oscillation of the reconstructed sequence. Namely, there was a 5~10 years dry-wet cycle in China. Moreover, as shown in Figure 7b, the amplitude during the periods AD 56~400, AD 1000~1400, and AD 1900~2000 was greater, indicating that these 3 periods of dry-wet change were more intense. The dominant multidecadal variation was 30~50 years (Figure 7c). The amplitude was greater than other periods during AD 800~1300. We analyzed the correlation between M-scPDSI (multidecadal component of our reconstructed scPDSI) and B2000 (the multidecadal component of B2000), considering lags of different lengths of time (0~50 years, Figure 8). The calculation found that when the M-scPDSI and B2000 have a time lag of 15~40 years, the correlation is better ( $r > 0.1$ ,  $p$ -value < 0.01). The positive correlation coefficient indicates that the dry-wet change of China has a 15~40 years lag response to the temperature change in the northern hemisphere on the multidecadal scale.



**Figure 8.** The histogram of correlation between M-scPDSI and B2000 with different lags lengths of time, M-scPDSI is short for the multidecadal component of our reconstructed scPDSI, and M-B2000 is short for the multidecadal component of B2000.

The centennial components of our reconstructed scPDSI were positively correlated with the centennial components of B2000 and DWJ2006, with correlation coefficients of 0.232 ( $n = 50$ ,  $p$ -value < 0.001) and 0.579 ( $n = 1283$ ,  $p$ -value < 0.001), respectively (Figure 7d). The frequency and amplitude of dry-wet change in China were basically consistent with the change of temperature in the northern hemisphere on the century scale, particularly during the MWP. The residual IMF was the remainder of our sequence from high frequency to low frequency, which reflected the long-term trend of our reconstructed sequence (Figure 7e). The long-term trend was broadly divided into two periods: a drying trend during the first

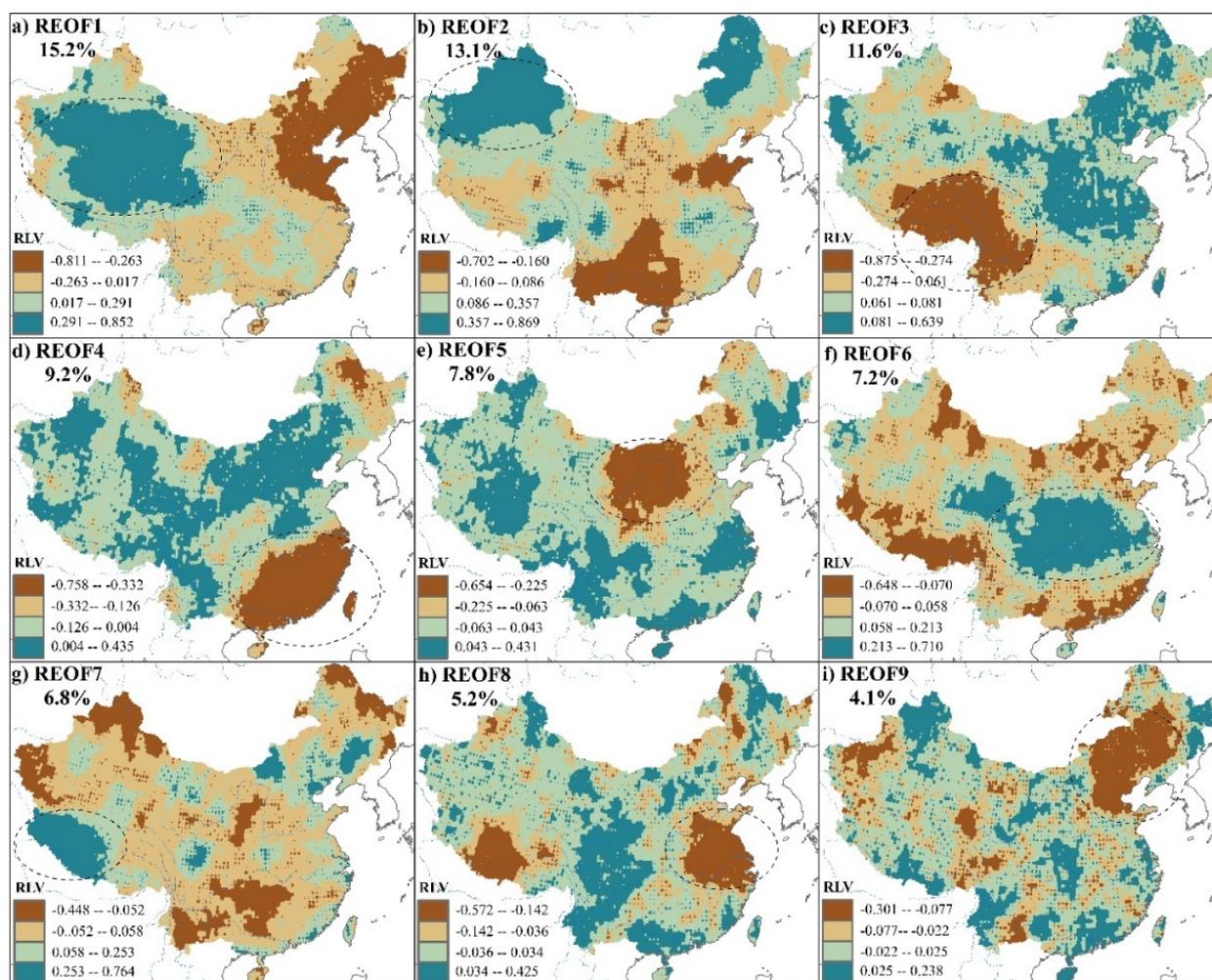
phase before the early 14th century, and then conditions gradually became wetter after a significant cold period in the Little Ice Age (LIA) until the beginning of the 20th century. This was consistent with the 4 groups of anomalous temperature sequences in the northern hemisphere (i.e., MJ2003, B2000, ECS2002, and DWJ2006), with correlation coefficients of 0.862, 0.469, 0.962, and 0.875, respectively. Moreover, during the MWP and LIA, the amplitude of the two significant warm and cold periods was greater than the other periods, indicating that the signal was stronger during these periods.

Overall, there were three different alternating fluctuation modes of dry-wet in China, which were all positively correlated with the amplitude and frequency of temperature in the northern hemisphere (in addition to the one-quarter-cycle delayed response over the multidecadal scale). Through comparison, it was found that the change characteristics of the reconstruction results of this study are basically the same as those of other research results. Therefore, we believe that the change characteristics of the reconstructed scPDSI in this study are credible.

#### 4.3.2. Spatial Characteristics

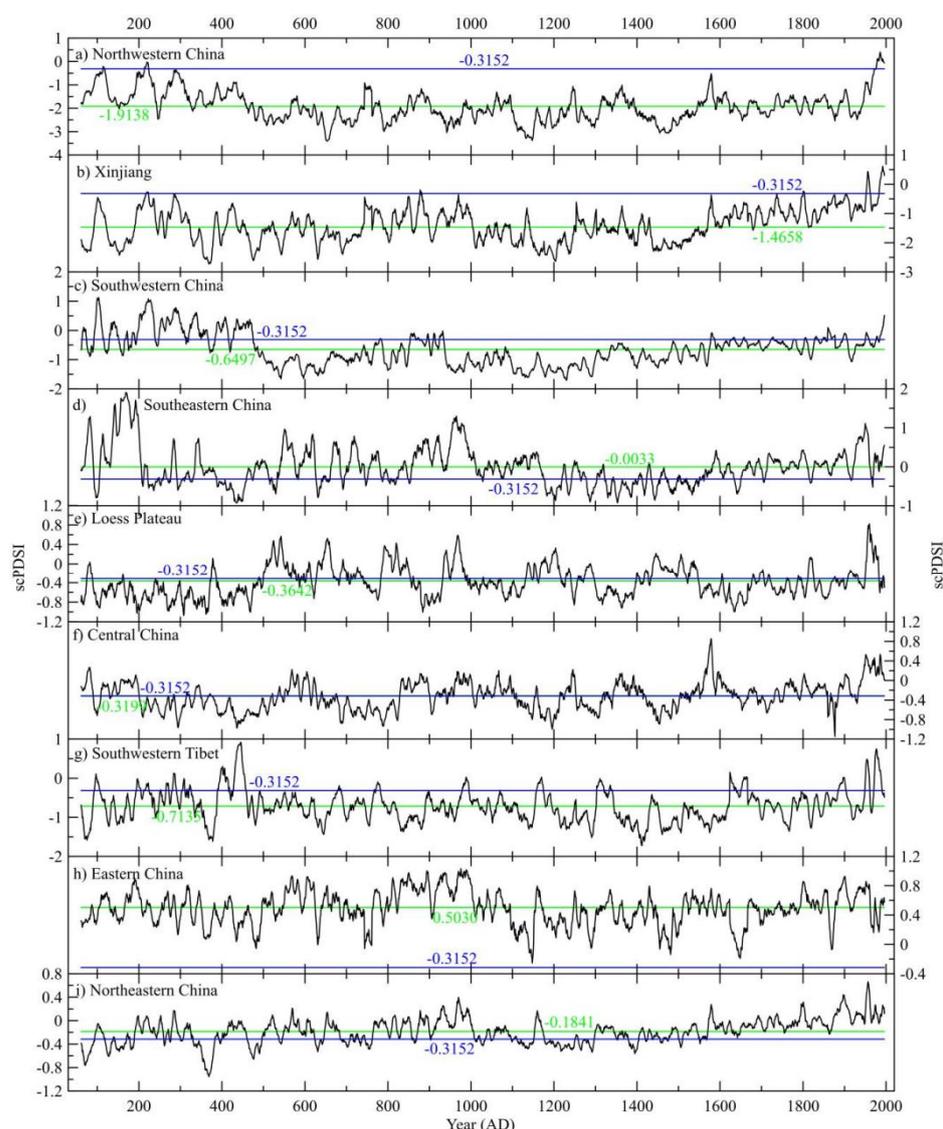
To explore the spatial distribution characteristics of the dry-wet change in China, our reconstructed scPDSI series (AD 56~2000) was decomposed by the REOF. The variance explained by the nine modes accounted for 15.2%, 13.1%, 11.6%, 9.2%, 7.8%, 7.2%, 6.8%, 5.2%, and 4.1% of the total variance, respectively. The cumulative explained variance contribution accounted for 80.2% of the total variance.

Figure 9 shows the spatial distribution characteristics of the first nine rotated loading vector (RLV) of the reconstructed scPDSI series, respectively. The first REOF (REOF1) leading mode of the reconstructed scPDSI series (Figure 9a) displayed the main loading in northwestern China (NWC), demonstrating an east-west anomalous dipole pattern (increased area covering southwestern China and a decreased area over northeastern China). The maximum RLV value of the load center was 0.852. The second REOF (REOF2) leading mode (Figure 9b) displayed main loading in Xinjiang (XJ). The maximum RLV value of the load center was 0.869. The third REOF (REOF3) leading mode (Figure 9c) displayed the main loading in southwestern China (SWC). The minimum RLV value of the load center was  $-0.875$ . The fourth REOF (REOF4) leading mode (Figure 9d) displayed the main loading in southeastern China (SEC). The minimum RLV value of the load center was  $-0.758$ . The fifth REOF (REOF5) leading mode (Figure 9e) displayed the main loading in the Loess Plateau (LP). The minimum RLV value of the load center was  $-0.654$ . The sixth REOF (REOF6) leading mode (Figure 9f) displayed the main loading in central China (CC). The maximum RLV value of the load center was 0.710. The seventh REOF (REOF7) leading mode (Figure 9g) displayed the main loading in southwestern Tibet (SWT). The maximum RLV value of the load center was 0.764. The eighth REOF (REOF8) leading mode (Figure 9h) displayed the main loading in eastern China (EC), illustrating a “sandwich” tri-pole pattern with an increased area covering the middle reaches of the Yellow River and Yangtze River, but decreasing over Tibet and eastern China. The minimum RLV value of the load center was  $-0.572$ . The ninth REOF (REOF9) leading mode (Figure 9i) displayed the main loading in northeastern China (NEC). The minimum RLV value of the load center was  $-0.301$ .



**Figure 9.** Spatial distribution characteristics of the rotated empirical orthogonal function of the reconstructed scPDSI series over the past two millennia (AD 56–2000).

To further analyze the history (AD 56~2000) of the dry-wet change in each characteristic region, we averaged the reconstructed scPDSI of the grid points contained in each area and then smoothed with the 10-year moving average (Figure 10). From the perspective of the mean value (AD 56~2000) of the reconstructed scPDSI of each sub-region, all regions were negative except EC (Figure 10h, mean value = 0.5030). This finding indicates that the whole country, except EC, is partially arid, and the arid situation of NWC (Figure 10a, mean value = -1.9138) was the most serious. Comparing the mean value of each sub-region with the mean value of all of China (Figure 6a, mean value = -0.3152), the mean value of SEC, EC, and NEC was higher than all of China. These three regions are China's humid zone, where the annual precipitation of EC and SEC was >800 mm, and the annual precipitation of NEC was >400 mm [34]. In addition, the mean value of CC and LP was almost equal to the mean value for all of China, indicating that the dry-wet change characteristics in CC and LP reflect the situation in all of China. NWC and XJ had the lowest mean value, indicating that all of NWC was the most severe drought region. This feature agrees with the actual situation in China. Therefore, our reconstructed scPDSI results were reasonable and reliable from the perspective of the characteristics of dry-wet change in each sub-region.



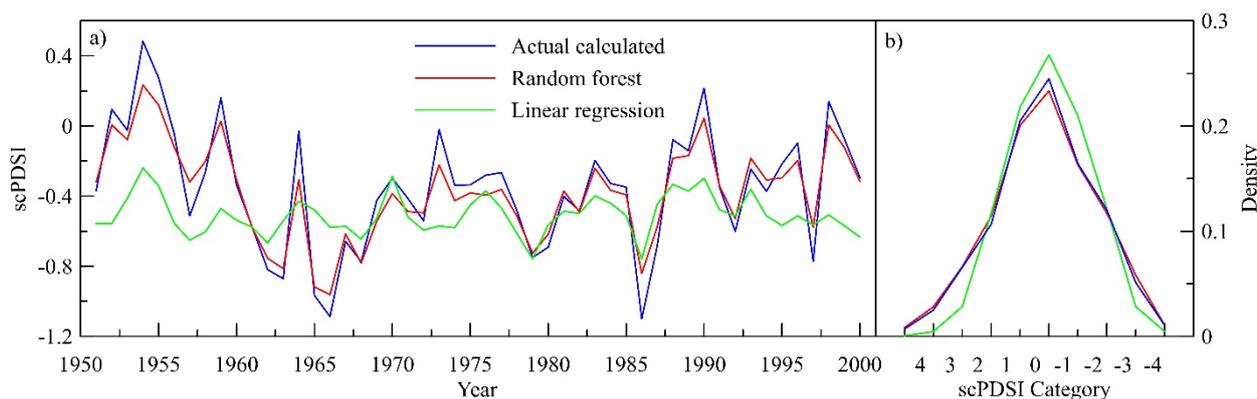
**Figure 10.** The dry-wet change of the reconstructed scPDSI of nine sub-regions. The green line represents the perennial average value of each sub-region, and the blue line represents the perennial average value of the whole of China.

## 5. Discussion

China has undergone 4 wet periods in 2000 years of history and 2 drying periods. The wet periods tended to be times to establish unity and a stable period for the ruling dynasty, while the drought period was a time of national turmoil and constant wars. We do not think this is a coincidence. As shown in the nine green areas in Figure 6a, a rapid droughts occurred during dynasty changes. This shows that climate change is an important precondition for the stability of the country. An agricultural society is particularly affected by the climate. Drought leads to a reduction in food production, but landlords and official taxes do not decrease, leaving many farmers without enough food to eat. The farmers embarked on a path of uprising, thus, pushing for change in the dynasty. The reconstruction results of this study were basically in accordance with historical facts based on identifying the significant drought/humid events, characterizing the wet and dry trends, and describing the turning points [74]. The results indicated that the reconstruction method was reasonable and that the results can be used as a reference for related research.

To estimate the effects of RF as applied to reconstruct the scPDSI, linear regression (LR) [75] was used to reconstruct the same scPDSI (3725 grid points,  $0.5^\circ \times 0.5^\circ$  resolution)

for the period of AD 1951~2000 and compare it with the actual calculated value (mentioned in Section 2.1, hereinafter referred to as AC), respectively. Figure 11a shows the variations in the average annual values between RF, LR, and AC, and Figure 11b shows the density distribution across different scPDSI categories. Several types of statistical factors were employed to evaluate the stability of different methods, including the mean, range, Pearson's correlation coefficient ( $r$ ), Nash-Sutcliffe efficiency (NSE) [76], and the percent bias (PBIAS, %) [77] (Table 4). Higher  $r$  and NSE values and lower PBIAS values indicate good performance of each reconstruction model. Namely, the closer the  $r$  and NSE values are to 1, the more accurate the model is. The closer the PBIAS value is to 0, the more accurate the model is.



**Figure 11.** Comparison of scPDSI between reconstructed results based on RF and LR (3725 grid points,  $0.5^\circ \times 0.5^\circ$  resolution) and actual calculated value (3725 grid points,  $0.5^\circ \times 0.5^\circ$  resolution). (a) Variation of annual average; and (b) density curve of different scPDSI categories.

**Table 4.** Several types of statistical factors of RF and LR.

	Mean	Max	Min	$r$	NSE	PBIAS
Actual calculated	−0.35	8.37	−6.98	/	/	/
Random forest	−0.37	7.55	−6.66	0.97	0.89	6.32%
Linear regression	−0.51	6.89	−5.98	0.60	0.07	44.79%

RF and LR were significantly positively correlated with AC (Table 4,  $n = 50$ ,  $p$ -value  $< 0.01$ ); that is, both RF and LR can be used for scPDSI reconstruction. However, it is clear from Figure 11a that the variations in RF and AC were basically the same, while LR fluctuated less, and the RF reconstruction results were more reliable and accurate than LR (Table 4). At the same time, the mean LR value was smaller, indicating that the LR reconstruction result was smaller (drier) than RF. As shown in Figure 11b, the density curve of different scPDSI categories of LR was the “lanky type”, while RF and AC were the “chunky type”, which shows that LR is less effective for characterizing extreme drought and flood events than RF. RF has significant advantages both in accuracy and stability compared to LR.

The RF has great application potential in data processing as an effective method for reconstructing drought variability because of advantages in computational accuracy and stability compared to the traditional method (LR). However, applying the algorithm to reconstruct scPDSI has crucial limitations in application. The reconstruction results of this study mainly show the two-dimensional spatial distribution of historical drought (more precisely, it is actually a simple collection of individual one-dimensional time series of regional averages of the drought index), but we failed to consider droughts as a (three-dimensional) space-time phenomenon [22]. Therefore, in future research, we plan to enrich the proxy data to improve the accuracy of the reconstruction model. At the same time, it will be necessary to detect atmospheric drought events that fully account for the dynamic space-time behavior of droughts more comprehensively and realistically.

## 6. Conclusions

The annual scPDSI of China during AD 56~2000 was reconstructed using multi-proxy records (tree-ring width data and stalagmite oxygen isotope  $\delta^{18}\text{O}$  data) based on RF. The temporal and spatial characteristics of the reconstruction results were analyzed and compared to previous studies. The data used in this study are reliable, the reconstruction method is accurate and feasible, and the reconstruction results are reasonable and credible. The four main conclusions of this study are as follows:

- (1) Based on the mean value of the reconstructed scPDSI (the mean value =  $-0.3151$ ), China has been in a drought-biased state for almost 2000 years.
- (2) There are three different alternating fluctuation modes (interannual scale, multi-decadal scale, and centennial scale) of dry-wet change, which were all positively correlated with the amplitude and frequency of the temperature in the northern hemisphere (in addition to the one-quarter-cycle delayed response over the multi-decadal scale).
- (3) China was divided into nine dry-wet change characteristic regions according to the characteristics of the different REOF leading mode distributions: northwestern China, Xinjiang, southwestern China, southeastern China, the Loess plateau, central China, southwestern Tibet, eastern China, and northeastern China. The dry-wet change characteristics in central China and the Loess plateau reflected the situation in all of China. Northwestern China was the most severe drought region.
- (4) The RF was highly accurate and stable for reconstructing drought variability in China compared with linear regression.

**Author Contributions:** B.Y. and X.C. conceived and designed the research. B.Y. performed data analysis and wrote the manuscript. Extensive editing and revisions to the final manuscript and conclusions were contributed by C.L., B.Y. and V.P.S. J.W. was responsible for collecting and preliminarily analyzing the data. All authors have read and agreed to the published version of the manuscript.

**Funding:** China (Grant No. U1911204,51861125203,91547202), National Key R&D Program of China (2021YFC3001000), the Chinese Academy of Engineering Consulting Project (2015-ZD-07-04-03), the Project for Creative Research from Guangdong Water Resources Department (Grant No. 2016-07, 2016-01), and the Research Program of Guangzhou Water Authority (2017).

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hao, Z.; Singh, V.P. Drought characterization from a multivariate perspective: A review. *J. Hydrol.* **2015**, *527*, 668–678. [[CrossRef](#)]
2. Wang, Z.; Li, J.; Lai, C.; Wang, R.; Chen, X.; Lian, Y. Drying tendency dominating the global major grain producing area. *Glob. Food Secur.* **2018**, *16*, 138–149. [[CrossRef](#)]
3. Lai, C.; Zhong, R.; Wang, Z.; Wu, X.; Chen, X.; Wang, P.; Lian, Y. Monitoring hydrological drought using long-term satellite-based precipitation data. *Sci. Total Environ.* **2019**, *649*, 1198–1208. [[CrossRef](#)] [[PubMed](#)]
4. Tjeldeman, E.; Barker, L.J.; Svoboda, M.D.; Stahl, K. Natural and human influences on the link between meteorological and hydrological drought indices for a large set of catchments in the contiguous United States. *Water Resour. Res.* **2018**, *54*, 6005–6023. [[CrossRef](#)]
5. Haslinger, K.; Hofstätter, M.; Kroisleitner, C.; Schöner, W.; Laaha, G.; Blöschl, G. Drivers of meteorological drought severity in the European Greater Alpine Region during the last two centuries. In Proceedings of the EGU General Assembly Conference Abstract, Vienna, Austria, 4–13 April 2018; Volume 20, p. 12075.
6. Kreyling, J.; Khan, M.A.A.; Sultana, F.; Babel, W.; Beierkuhnlein, C.; Foken, T.; Jentsch, A. Drought effects in climate change manipulation experiments: Quantifying the influence of ambient weather conditions and rain-out shelter artifacts. *Ecosystems* **2017**, *20*, 301–315. [[CrossRef](#)]
7. Xu, K.; Yang, D.; Yang, H.; Li, Z.; Qin, Y.; Shen, Y. Spatio-temporal variation of drought in China during 1961~2012: A climatic perspective. *J. Hydrol.* **2015**, *526*, 253–264. [[CrossRef](#)]
8. Ge, Y.; Apurv, T.; Cai, X. Spatial and temporal patterns of drought in the Continental US during the past century. *Geophys. Res. Lett.* **2016**, *43*, 6294–6303. [[CrossRef](#)]
9. Carbone, G.J.; Lu, J.; Brunetti, M. Estimating uncertainty associated with the standardized precipitation index. *Int. J. Climatol.* **2018**, *38*, e607–e616. [[CrossRef](#)]

10. McKee, T.B.; Doesken, N.J.; Kleist, J. The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology, Anaheim, CA, USA, 17–22 January 1993; Volume 17, pp. 179–183.
11. Vicente-Serrano, S.M.; Beguería, S.; López-Moreno, J.I. A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Clim.* **2010**, *23*, 1696–1718. [[CrossRef](#)]
12. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
13. Palmer, W.C. Meteorological drought. In *US Department of Commerce; Weather Bureau: Washington, DC, USA, 1965.*
14. Alley, W.M. The Palmer Drought Severity Index: Limitations and Assumptions. *J. Appl. Meteorol. Climatol.* **1984**, *23*, 1100–1109. [[CrossRef](#)]
15. Wells, N.; Goddard, S.; Hayes, M.J. A self-calibrating Palmer drought severity index. *J. Clim.* **2004**, *17*, 2335–2351. [[CrossRef](#)]
16. Dai, A.; Trenberth, K.E.; Qian, T. A global data set of palmer drought severity index for 1870–2002: Relationship with soil moisture and effects of surface warming. *J. Hydrometeorol.* **2009**, *5*, 1117–1130. [[CrossRef](#)]
17. Van der Schrier, G.; Barichivich, J.; Briffa, K.R.; Jones, P.D. A scPDSI-based global data set of dry and wet spells for 1901–2009. *J. Geophys. Res. Atmos.* **2013**, *118*, 4025–4048. [[CrossRef](#)]
18. Liu, L.; Hong, Y.; Bednarczyk, C.N.; Yong, B.; Shafer, M.A.; Riley, R.; Hocker, J.E. Hydroclimatological drought analyses and projections using meteorological and hydrological drought indices: A case study in Blue River Basin, Oklahoma. *Water Resour. Manag.* **2012**, *26*, 2761–2779. [[CrossRef](#)]
19. Wang, Z.; Lai, C.; Chen, X.; Yang, B.; Zhao, S.; Bai, X. Flood hazard risk assessment model based on random forest. *J. Hydrol.* **2015**, *527*, 1130–1141. [[CrossRef](#)]
20. Burke, E.J.; Brown, S.J.; Christidis, N. Modeling the recent evolution of global drought and projections for the twenty-first century with the Hadley Centre climate model. *J. Hydrometeorol.* **2006**, *7*, 1113–1125. [[CrossRef](#)]
21. Wang, Z.; Li, J.; Lai, C.; Zeng, Z.; Zhong, R.; Chen, X.; Wang, M. Does drought in China show a significant decreasing trend from 1961 to 2009? *Sci. Total Environ.* **2017**, *579*, 314–324. [[CrossRef](#)]
22. Haslinger, K.; Blöschl, G. Space-time patterns of meteorological drought events in the European Greater Alpine Region over the past 210 years. *Water Resour. Res.* **2017**, *53*, 9807–9823. [[CrossRef](#)]
23. Woodhouse, C.A.; Overpeck, J.T. 2000 Years of Drought Variability in the Central United States. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 2693–2714. [[CrossRef](#)]
24. Yang, B.; Qin, C.; Wang, J. A 3500-year tree-ring record of annual precipitation on the northeastern Tibetan Plateau. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2903–2908. [[CrossRef](#)] [[PubMed](#)]
25. Cai, Q.; Liu, Y.; Lei, Y.; Bao, G.; Sun, B. Reconstruction of the March–August PDSI since 1703 AD based on tree rings of Chinese pine (*Pinus tabulaeformis* Carr.) in the Lingkong Mountain, southeast Chinese loess Plateau. *Clim. Past* **2014**, *9*, 6311–6344. [[CrossRef](#)]
26. Linderholm, H.W.; Björklund, J.A.; Seftigen, K.; Gunnarson, B.E.; Grudd, H.; Jeong, J.H.; Drobyshev, I.; Liu, Y. Dendroclimatology in Fennoscandia—from past accomplishments to future potential. *Clim. Past* **2010**, *6*, 93–114. [[CrossRef](#)]
27. Büntgen, U.; Tegel, W.; Nicolussi, K.; McCormick, M.; Frank, D.; Trouet, V.; Kaplan, J.O.; Herzig, F.; Heussner, K.-U.; Wanner, H.; et al. 2500 years of European climate variability and human susceptibility. *Science* **2011**, *331*, 578–582. [[CrossRef](#)]
28. Liu, Y.; Zhang, X.; Song, H.; Cai, Q.; Li, Q.; Zhao, B.; Mei, R. Tree-ring-width-based PDSI reconstruction for central Inner Mongolia, China over the past 333 years. *Clim. Dyn.* **2017**, *48*, 867–879. [[CrossRef](#)]
29. Gaire, N.; Dhakal, Y.R.; Shah, S.K. Drought (scPDSI) reconstruction of trans-Himalayan region of central Himalaya using *Pinus wallichiana* tree-rings. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2019**, *514*, 251–264. [[CrossRef](#)]
30. Esper, J.; Frank, D.; Büntgen, U.; Verstege, A.; Luterbacher, J.; Xoplaki, E. Long-term drought severity variations in Morocco. *Geophys. Res. Lett.* **2007**, *34*, 244. [[CrossRef](#)]
31. Cook, E.R.; Anchukaitis, K.J.; Buckley, B.M.; D’Arrigo, R.D.; Jacoby, G.C.; Wright, W.E. Asian monsoon failure and megadrought during the last millennium. *Science* **2010**, *328*, 486–489. [[CrossRef](#)]
32. Cook, E.R.; Seager, R.; Kushnir, Y.; Briffa, K.R.; Büntgen, U.; Frank, D.; Baillie, M. Old World megadroughts and pluvials during the Common Era. *Sci. Adv.* **2015**, *1*, e1500561. [[CrossRef](#)]
33. Palmer, J.G.; Cook, E.R.; Turney, C.S.; Allen, K.; Fenwick, P.; Cook, B.I.; O’Donnell, A.; Lough, J.; Grierson, P.; Baker, P. Drought variability in the eastern Australia and New Zealand summer drought atlas (ANZDA, CE1500–2012) modulated by the Interdecadal Pacific Oscillation. *Environ. Res. Lett.* **2015**, *10*, 124002. [[CrossRef](#)]
34. Shi, F.; Zhao, S.; Guo, Z.; Goosse, H.; Yin, Q. Multi-proxy reconstructions of May–September precipitation field in China over the past 500 years. *Clim. Past* **2017**, *13*, 1919–1938. [[CrossRef](#)]
35. He, M.; Yang, B.; Datsenko, N.M. A six hundred-year annual minimum temperature history for the central Tibetan Plateau derived from tree-ring width series. *Clim. Dyn.* **2014**, *43*, 641–655. [[CrossRef](#)]
36. Nguyen, H.T.; Galelli, S. A linear dynamical systems approach to streamflow reconstruction reveals history of regime shifts in northern Thailand. *Water Resour. Res.* **2018**, *54*, 2057–2077. [[CrossRef](#)]
37. Littell, J.S.; Pederson, G.T.; Gray, S.T.; Tjoelker, M.; Hamlet, A.F.; Woodhouse, C.A. Reconstructions of Columbia River streamflow from tree-ring chronologies in the Pacific Northwest, USA. *J. Am. Water Resour. Assoc.* **2016**, *52*, 1121–1141. [[CrossRef](#)]
38. Allen, K.J.; Nichols, S.C.; Evans, R.; Allie, S.; Carson, G.; Ling, F.; Baker, P.J. A 277 year cool season dam inflow reconstruction for Tasmania, southeastern Australia. *Water Resour. Res.* **2017**, *53*, 400–414. [[CrossRef](#)]

39. Shi, F.; Ge, Q.; Yang, B.; Li, J.; Yang, F.; Ljungqvist, F.C.; Xu, C. A multi-proxy reconstruction of spatial and temporal variations in Asian summer temperatures over the last millennium. *Clim. Chang.* **2015**, *131*, 663–676. [[CrossRef](#)]
40. Yang, F.; Wang, N.; Shi, F.; Ljungqvist, F.C.; Zhao, S.; Liu, T. The spatial distribution of precipitation over the West Qinling region, China, AD 1470–2000. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **2016**, *443*, 278–285. [[CrossRef](#)]
41. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
42. Wang, H.; Chen, Y.; Pan, Y.; Li, W. Spatial and temporal variability of drought in the arid region of China and its relationships to teleconnection indices. *J. Hydrol.* **2015**, *523*, 283–296. [[CrossRef](#)]
43. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
44. Chen, H.; Liang, Z.; Liu, Y.; Jiang, Q.; Xie, S. Effects of drought and flood on crop production in China across 1949–2015: Spatial heterogeneity analysis with Bayesian hierarchical modeling. *Nat. Hazards* **2018**, *92*, 525–541. [[CrossRef](#)]
45. Shao, D.; Chen, S.; Tan, X.; Gu, W. Drought characteristics over China during 1980–2015. *Int. J. Climatol.* **2018**, *38*, 3532–3545. [[CrossRef](#)]
46. Wang, Z.; Zhong, R.; Lai, C.; Zeng, Z.; Lian, Y.; Bai, X. Climate change enhances the severity and variability of drought in the Pearl River Basin in South China in the 21st century. *Agric. For. Meteorol.* **2018**, *249*, 149–162. [[CrossRef](#)]
47. Zhao, A.; Zhang, A.; Cao, S.; Liu, X.; Li, J.; Cheng, D. Responses of vegetation productivity to multi-scale drought in Loess Plateau, China. *Catena* **2018**, *163*, 165–171. [[CrossRef](#)]
48. Zhang, L.; Wu, P.; Zhou, T.; Xiao, C. ENSO Transition from La Niña to El Niño Drives Prolonged Spring-Summer Drought over North China. *J. Clim.* **2018**, *31*, 3509–3523. [[CrossRef](#)]
49. Yu, M.; Li, Q.; Hayes, M.J.; Svoboda, M.D.; Heim, R.R. Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010? *Int. J. Climatol.* **2014**, *34*, 545–558. [[CrossRef](#)]
50. Fang, K.; Gou, X.; Chen, F.; Liu, C.; Davi, N.; Li, J.; Li, Y. Tree-ring based reconstruction of drought variability (1615–2009) in the Kongtong Mountain area, northern China. *Glob. Planet. Change* **2012**, *80*, 190–197. [[CrossRef](#)]
51. Cai, Q.; Liu, Y. Climatic response of Chinese pine and PDSI variability in the middle Taihang Mountains, north China since 1873. *Trees* **2013**, *27*, 419–427. [[CrossRef](#)]
52. Li, J.; Shi, J.; Zhang, D.D.; Yang, B.; Fang, K.; Yue, P.H. Moisture increase in response to high-altitude warming evidenced by tree-rings on the southeastern Tibetan Plateau. *Clim. Dyn.* **2017**, *48*, 649–660. [[CrossRef](#)]
53. Su, J.; Gou, X.; Deng, Y.; Zhang, R.; Liu, W.; Zhang, F.; Zheng, W. Tree growth response of *Fokienia hodginsii* to recent climate warming and drought in southwest China. *Int. J. Biometeorol.* **2017**, *61*, 2085–2096. [[CrossRef](#)]
54. Gou, X.; Deng, Y.; Gao, L.; Chen, F.; Cook, E.; Yang, M.; Zhang, F. Millennium tree-ring reconstruction of drought variability in the eastern Qilian Mountains, northwest China. *Clim. Dyn.* **2015**, *45*, 1761–1770. [[CrossRef](#)]
55. Yang, B.; Kang, S.; Ljungqvist, F.C.; He, M.; Zhao, Y.; Qin, C. Drought variability at the northern fringe of the Asian summer monsoon region over the past millennia. *Clim. Dyn.* **2014**, *43*, 845–859. [[CrossRef](#)]
56. Yang, H.; Yang, D.; Hu, Q.; Lv, H. Spatial variability of the trends in climatic variables across China during 1961–2010. *Theor. Appl. Climatol.* **2015**, *120*, 773–783. [[CrossRef](#)]
57. International Tree Ring Data Bank. Available online: <http://www.ncdc.noaa.gov/paleo/treering.html> (accessed on 7 January 2022).
58. Wang, Y.; Cheng, H.; Edwards, R.L.; He, Y.; Kong, X.; An, Z.; Li, X. The Holocene Asian monsoon: Links to solar changes and North Atlantic climate. *Science* **2005**, *308*, 854–857. [[CrossRef](#)]
59. Tan, M.; Liu, T.; Hou, J.; Qin, X.; Zhang, H.; Li, T. Cyclic rapid warming on centennial-scale revealed by a 2650-year stalagmite record of warm season temperature. *Geophys. Res. Lett.* **2003**, *30*, 31. [[CrossRef](#)]
60. Adusumilli, S.; Bhatt, D.; Wang, H.; Devabhaktuni, V.; Bhattacharya, P. A novel hybrid approach utilizing principal component regression and random forest regression to bridge the period of GPS outages. *Neurocomputing* **2015**, *166*, 185–192. [[CrossRef](#)]
61. Zhang, D. Test of calibration on the paleoclimatic proxy data by using Chinese historical records. *Adv. Clim. Change Res.* **2010**, *6*, 70–72. [[CrossRef](#)]
62. Allen, E.B.; Rittenour, T.M.; DeRose, R.J.; Bekker, M.F.; Kjelgren, R.; Buckley, B.M. A tree-ring based reconstruction of Logan River streamflow, northern Utah. *Water Resour. Res.* **2013**, *49*, 8579–8588. [[CrossRef](#)]
63. Cook, E.R.; Meko, D.M.; Stahle, D.W.; Cleaveland, M.K. Drought reconstructions for the continental United States. *J. Clim.* **1999**, *12*, 1145–1162. [[CrossRef](#)]
64. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
65. Lorenz, E.N. *Empirical Orthogonal Functions and Statistical Weather Prediction*; Science Report No. 1; MIT: Cambridge, MA, USA, 1956; pp. 1–49.
66. Richman, M.B. Review article, rotation of principal components. *J. Climatol.* **1986**, *6*, 293–355. [[CrossRef](#)]
67. Liu, Z.; Zhou, P.; Zhang, F.; Liu, X.; Chen, G. Spatiotemporal characteristics of dryness/wetness conditions across Qinghai Province, Northwest China. *Agric. For. Meteorol.* **2003**, *182*, 101–108. [[CrossRef](#)]
68. Ashok, K.; Behera, S.K.; Rao, S.A.; Weng, H.; Yamagata, T. El Niño Modoki and its possible teleconnection. *J. Geophys. Res. Ocean.* **2007**, *112*, 3798. [[CrossRef](#)]
69. Mann, M.E.; Jones, P.D. Global surface temperatures over the past two millennia. *Geophys. Res. Lett.* **2003**, *30*, 17814. [[CrossRef](#)]

70. Esper, J.; Cook, E.R.; Schweingruber, F.H. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* **2002**, *295*, 2250–2253. [[CrossRef](#)]
71. Cook, E.R.; Esper, J.; D'Arrigo, R.D. Extra-tropical Northern Hemisphere land temperature variability over the past 1000 years. *Quat. Sci. Rev.* **2004**, *23*, 2063–2074. [[CrossRef](#)]
72. D'Arrigo, R.; Wilson, R.; Jacoby, G. On the long-term context for late twentieth century warming. *J. Geophys. Res. Atmos.* **2006**, *111*, 6352. [[CrossRef](#)]
73. Briffa, K.R. Annual climate variability in the Holocene: Interpreting the message of ancient trees. *Quat. Sci. Rev.* **2000**, *19*, 87–105. [[CrossRef](#)]
74. Zhang, D.E.; Liang, Y.Y. A long lasting and Extensive Drought Event over China during 1876–1878. *Adv. Clim. Chang. Res.* **2010**, *6*, 106–112.
75. Kass, R.E. Nonlinear regression analysis and its applications. *J. Am. Stat. Assoc.* **1990**, *85*, 594–596. [[CrossRef](#)]
76. McCuen, R.H.; Knight, Z.; Cutter, A.G. Cutter. Evaluation of the Nash-Sutcliffe efficiency index. *J. Hydrol. Eng.* **2006**, *11*, 597–602. [[CrossRef](#)]
77. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [[CrossRef](#)]